Characterizing the Cu-only SOD repeat protein of the zebrafish *Danio rerio*

by
Mengrui Wang

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science

Baltimore, Maryland
April, 2019

Abstract

The Cu-containing superoxide dismutases (SODs) are enzymes which participate in the metabolism of reactive oxygen species (ROS) by catalyzing the dismutation of superoxide free radical into oxygen and hydrogen peroxide. Catalysis is driven by a Cu cofactor and, in most cases, a bridging Zn that enhances Cu catalysis and stabilizes the protein. Such Cu/Zn SODs are widespread; however, the Culotta lab has recently found a new class of SODs which functions without a Zn cofactor. These Cu-only enzymes are widespread in fungi and oomycetes as single subunit enzymes. Interestingly, they are also found in certain classes of the animal kingdom, appearing as tandem repeats of Cu-only SOD-like protein domains in large polypeptides we refer to as CSRP (Cu-only SOD repeat proteins).

In this work, we specifically studied CSRP in zebrafish. The main goals were (i) to test whether the recombinant zebrafish CSRP domains can also function as Cu-binding SOD proteins to remove superoxide, and (ii) to ascertain whether the CSRP protein is expressed in the fish as a full-length protein containing multiple tandem repeats of a SOD-like domain. We found that an isolated zebrafish CSRP domain produced recombinantly in *E. coli* tends to oligomerize, unlike the fungal Cu-only SOD, which is always a monomer. Using equilibrium dialysis, we found that an isolated CSRP domain from zebrafish can bind Cu, and this is the first study to suggest that it has SOD activity. In the future, we will further examine whether the full-length recombinant CSRP can act as a SOD, or whether GPI-anchored CSRP in live cells also has SOD activity.

We also examined the expression of CSRP in zebrafish tissues and found that CSRP was most abundantly expressed the in heart and fin. Interestingly, the protein appears to be expressed as dimers and trimers, suggesting that the full-length protein is processed into various units of Cu-

only SOD domains. The high expression of CSRP in heart and fin is of note because in zebrafish, these two tissues are capable of regeneration involving bursts of reactive oxygen species. In the future we can study whether CSRP has evolved to help deal with reactive oxygen species generated during tissue regeneration, wound healing and development.

Advisor

Dr. Valeria Culotta, Ph.D., Department of Biochemistry and Molecular Biology

Thesis Reader

Dr. Scott Bailey, Ph.D., Department of Biochemistry and Molecular Biology

Acknowledgements

I would like to express my gratitude and thanks to my thesis advisor, Dr. Valeria Culotta for her guidance and support throughout the course of this study.

I would also like to give special thanks to Dr. Scott Bailey for his invaluable feedbacks to better this work.

Lastly, I would like to express special thanks to my colleague, Natalie Robinett for her support and company during the whole process. I would also like to give great thanks to my other colleagues in Culotta Lab: Sabrina Schatzman, Angelique Besold, Ed Culbertson and Griffin Kowalewski for their tremendous help.

Table of Contents

List of Figures

1. INTRODUCTION

1.1 Cu-containing superoxide dismutases

Superoxide dismutase was first discovered in 1969 by McCord and Fridovich as a Cu metalloprotein that can catalyze the dismutation of superoxide free radicals via a ping-pong reaction mechanism [1] (2 equations below).

$$Cu\ (II) + O_2^{\bullet -} \rightarrow Cu\ (I) + O_2$$

$$Cu\ (I) + O_2^{\bullet -} + 2H^+ \rightarrow Cu\ (II) + H_2O_2$$

Zn was later detected in this cuproprotein, and this Cu and Zn bimetalloenzyme was defined as the mammalian SOD1 [2]. Since the discovery of Zn as a cofactor in SODs, other classes of SODs that use Fe, Mn or Ni as cofactors have been discovered. Although these SODs have different primary sequences and structures, they all use a redox active metal cofactor to catalyze the dismutation of superoxide [3]. SODs participate in cell oxidative stress protection, particularly by dismutating superoxide generated during metabolism [4, 5], and also function in cell signaling through the local production of H_2O_2 [6-8]. In addition, many SODs can work as virulence factors and are able to defend the oxidative burst which happens at the host-pathogen interface [9, 10].

1.2 The ubiquitous Cu/Zn SODs

The best-studied class of SOD enzymes is that of the bimetallic Cu/Zn SODs, which are widely conserved from bacteria to mammals and exist in both intracellular and extracellular sites. All eukaryotes contain an intracellular Cu/Zn SOD called SOD1 [1-3, 11, 12]. SOD1 plays a role in superoxide-caused oxidation stress protection, including that during mitochondrial respiratory chain activity, and also participates in cell signaling involving H_2O_2 generation and peroxide-sensitive kinase signaling [6, 7]. Interestingly, the action of SOD1 varies in different hosts. For

example, the Culotta lab found that in *Saccharomyces cerevisiae* yeast, SOD1 generates H_2O_2 to stabilize the casein kinase YCK1 for glucose sensing, while in *Candida albicans* yeast, SOD1 works in the glucose repression pathway but plays no role in casein kinase signaling [7, 13]. Although SOD1 primarily exists in the cytosol [14], it is also found in other intracellular compartments, such as the mitochondrial intermembrane space (IMS) [15-17], the secretory pathway [18, 19] and even the nucleus [20, 21]. The Culotta lab previously demonstrated that in the IMS of mitochondria, a fraction of active SOD1 localizes together with its Cu chaperone, CCS, to help protect against reactive oxygen species (ROS) [17]. In the nucleus, SOD1 functions in bakers' yeast to control gene expression during gene responses to DNA damage and Cu starvation stress [20, 21].

Cu/Zn SODs also exist in extracellular locations, as does their superoxide substrate, since superoxide cannot cross a biological membrane. Certain bacteria express Cu/Zn SODs in their periplasmic space [22-24], and in pathogenic bacteria, these SODs protect against the oxidative stress from the host immune system [9, 10]. Many eukaryotes also express another extracellular Cu/Zn SOD called ecSOD, which was first reported by Marklund in 1982 [25] and is a secreted protein in extracellular sites [3, 25-28]. The superoxide substrate for ecSOD originates from cell surface NADPH oxidase (NOX) enzymes [29], which make extracellular superoxide to combat pathogens and transmit signals involving reactive oxygen and nitrogen species [8].

Cu/Zn SODs can be multimeric, such as homodimeric SOD1 and tetrameric ecSOD [3, 30]. Each monomer contains several key features: a Greek key β-barrel fold, Cu and Zn sites, a conserved disulfide, an active site arginine, and an extended loop VII, which is also called the electrostatic loop (ESL) in eukaryotes [31, 32]. The catalytic Cu (II) is bound to a H_2O molecule and four histidines, one of which, His-63, also binds Zn (Fig. 1A). His-63 is thus called a "dynamic" histidine. With the

oxidation of superoxide, Cu (II) is reduced to Cu(I) and ceases to bind to His-63. The Zn cofactor, however, remains bridged to the His-63 during catalysis [31, 33] (Fig. 1A).

The Zn cofactor helps to enhance re-oxidation of Cu (I) to Cu (II) during catalysis [34, 35]. It also facilitates the stabilization of the protein structure [35-37]. Additional necessary features of Cu/Zn SODs include a disulfide [38] and an active site arginine that can stabilize the superoxide substrate [39] (Fig. 1A). The ESL contains four amino acid residues which are involved in electrostatic interactions to help guide the superoxide into the active site; the ESL also stabilizes Cu binding [39-43].

1.3 Cu-only SODs in the fungal kingdom

Extracellular SODs in fungal pathogens, like those in bacteria, work as a first line to defend against superoxide [9]. Extracellular SOD5 from a fungal pathogen, *Candida albicans,* is the most commonly studied example [44] (Fig. 2B). *C. albicans* expresses three extracellular SODs (SOD4, SOD5, SOD6) that are attached to the cell wall using GPI anchors [45]. Of the three SODs, SOD5 expression is highest in various forms of candidiasis [46-51]. For many years, these three extracellular SODs were considered bimetallic Cu/Zn SODs [44, 52], but in 2014, the Culotta lab found that they represent a new class of Cu-only SODs [53]. Structural studies and biochemical analyses of *C. albicans* SOD5 demonstrated that SOD5 lacks two Zn-binding histidines and also lacks the ESL [53], where an open active site is formed to make Cu more solvent exposed (Fig. 1B). Experiments attempting to bind SOD5 with Zn failed, which means that SOD5 functions without the Zn as a cofactor [53]. Despite its structural difference from Cu/Zn SODs, Cu-only SOD5 is highly efficient; its rate of superoxide catalysis is close to diffusion limits, k_{cat}=1.8 x 10^9 $M^{-1}s^{-1}$ at pH 6.0 [53, 54].

As previously mentioned, Zn promotes catalysis of Cu/Zn SODs by interacting with the dynamic His-63 [34, 35, 55]. Here in Cu-only SOD5, the dynamic histidine bridges to a conserved Glu-110 (Fig. 1B), which appears to play a role similar to Zn [54].

The highly charged ESL was previously considered to be essential for directing the superoxide substrate to the active site in Cu/Zn SODs [39, 41, 42]. In Cu-only SOD5, although the ESL is absent, SOD5 catalysis is still affected by ionic strength, which suggests that the SOD may have another charged region for substrate guidance [53]. In Cu/Zn SOD1, the ESL also functions to stabilize the Cu-binding site through interactions between ESL Asp-124 and the Cu- coordinating His-46 [34, 37]. In Cu-only SOD5, Asp-113 in the active site may substitute for ESL Asp-124 through a hydrogen bond network to the Cu-coordinating His-75 (Fig. 1B). Interestingly, Asp-113 exists in all Cu-containing SODs, but with different functions; in Cu/Zn SODs, Asp-113 functions by coordinating Zn, while in Cu-only SODs, it interacts with the Cu site through hydrogen bonding to stabilize Cu binding [54]. This specific usage of Asp-113 together with Glu-110 may explain how eukaryotic Cu-only SODs work without Zn and the ESL.

It should be noted that the aforementioned features of Glu-110 and Asp-113 may not be applicable to prokaryotic Cu-only SODs. For example, *Mycobacterial* (Mt) SodC lacks an equivalent of SOD5 Glu-110; instead, an aspartate plays a similar role with the histidine [54, 56]. Secondly, Mt SodC contains an extended loop VII, or the equivalent to the ESL in eukaryotic Cu/Zn SODs, and the Cu site is closed rather than open. Therefore, Mt SodC can be considered a mixture of Cu/Zn and fungal Cu-only SODs [32].

Upon examining the metalation process, the advantage of expressing a Cu-only SOD versus a Cu/Zn SOD becomes clear. In mammals, Cu and Zn are loaded to the ecSOD while the protein

moves through the secretory pathway, and then ecSOD reaches the cell surface in an active form [57, 58]. In comparison, the Cu-only SOD5 is secreted in *C. albicans* without Cu and is activated upon the addition of environmental Cu [53]. Without a Zn site, mis-metalation events, such as Cu moving to the Zn site, [55] are avoided. In addition, the Cu site of *C. albicans* SOD5 (Fig. 2B) appears resistant to mis-metalation even if the environmental Zn fluctuates [53], so as to retain intact enzyme activity.

1.4 SOD5-like protein domains in animals

We define a SOD5-like protein as one with the key features mentioned above, which lacks Zn sites and the ESL but keeps Glu-110 and Asp-113 equivalents. When searching kingdoms other than the fungal, the only organisms we found which express around 20-30 kDa SOD5-like SODs are oomycetes (Fig. 2A). Similar to those in fungi, the oomycete proteins are also secreted, GPI-anchored extracellular Cu-only SODs [54] which are thought to result from a horizontal gene transfer event from a fungal ancestor [59, 60].

SOD5-like proteins also exist in specific classes of animals, but not in any identified plants, protists, archaea or eubacteria (Fig. 2A). Currently, we hypothesize that approximately 1.5 billion years ago, just before the split of animals and fungi, a primordial Cu/Zn SOD lost its Zn site and the ESL loop. This SOD then evolved in fungi as a monomeric extracellular Cu-only SOD. However, in animals, the gene was apparently duplicated twice to give four repeats of a Cu-only SOD on a single large polypeptide of ≥100 kDa. The animal proteins are thought to be outside of the cell linked by GPI anchors, as in fungi and oomycetes. They also exhibit the structural features of SOD5-like SODs including Cu-only, lacking the ESL and Zn-binding sequences with a predicted open active site. We define animal SOD repeat proteins as CSRP (Cu-only SOD repeat proteins) [32].

One example is the CSRP[1] of the zebrafish *Danio rerio.* As seen in Fig. 3A, this protein contains four tandem repeats of a SOD5-like domain. Each domain is predicted to contain a SOD5-like Greek key beta-barrel fold and other features such as disulfide cysteines, an active site arginine and active site residues Glu-110 and Asp-113, with the exception of domain four, in which Glu-110 is replaced by methionine [32]. All except the fourth domain also have four Cu-binding histidines, while the ESL and Zn site are missing in all four domains. The reason is unknown why the fourth one has no Cu site (Fig. 3B and C). This identical pattern of three Cu-binding repeats plus a fourth non-Cu binding domain is conserved in the class of bony fishes/Osteichthyes, such as red piranha, common carp and Atlantic salmon [32].

One of the most interesting features of CSRP is its distribution among animals. CSRPs exist throughout the animal kingdom from the unicellular *Capsaspora owczarzaki* to vertebrate teleosts. But there is no evidence for the existence of CSRP in reptiles, birds or mammals [32]. The reason for this curious distribution is unknown, but it is worth noting that all of the animals without CSRP are land animals with lungs, and have a low capacity for tissue regeneration. As described in more detail in this work, tissue regeneration in CSRP-expressing classes of animals, including, fish, insects and amphibians, proceeds through a large burst in ROS production [61-64].

Depending on the class of animals, the pattern of the putative Cu-binding repeats varies. For example, in the Pacific oyster *Crassostrea gigas*, all four repeats are "perfect" in that they retain all the features of Cu-only SODs, including an intact Cu site. In flying insects, only one of the four domains (domain 3) is predicted to retain Cu-binding capacity. To date, these curious SOD-like

[1] while this thesis was in its final stages of preparation, the zebrafish genome database or Zfin requested we assign the name *CUSR* for si:dkey-16m19.2 rather than *CSRP*, as the name *CSRP* is already assigned to another protein in zebrafish *D. rario.*

proteins have been mis-labeled as Cu/Zn SODs, including the description of CSRP from *Anopholes gambiae* [65], the Pacific oyster *Crassostrea gigas* [66], and the moth *Bombyx mori* [67]. These genes are known to be expressed, for example, in the mantle-gonad tissue in *Crassostrea gigas* [66]. The proteins are also present in placozoans, as has been shown in proteomic analysis [68]. One of the most recent studies has specifically searched for SOD in the *Bombyx mori* genome. *Bombyx mori* is a lepidopteran insect whose body size is large enough for gene expression to be easily examined at the tissue level. In addition to the classical Cu/Zn SOD1, mitochondrial Mn-containing SOD2, and Cu/Zn ecSOD3, the group reports three additional "Cu/Zn" SODs: BmSOD4, BmSOD5, and BmSOD6, of which BmSOD6 is equivalent to CSRP of this organism. From this study, it is known that BmSOD6 is mainly expressed in the testes and ovaries, but also expressed in other tissues, such as the fat bodies and midguts [67].

1.5 Zebrafish *Danio rerio* as a model organism to study CSRP

What is CSRP? Does this protein function as a multimer of extracellular Cu-only SODs? To date, there has been no biochemical analysis of these proteins in any organism. Second, is this protein expressed as a large, full-length protein with repeating SOD enzyme units, or are these clipped off to produce monomers, or dimers/trimers of SOD-like proteins? To address these questions, we have chosen the zebrafish model, *Danio rerio,* for this study.

The zebrafish *Danio rerio* has been used as model organism in research since the 1960s. It is now one of the major vertebrate model used for genetics to study disease because of the many characteristics it possesses and the many technical advantages it has when compared to other vertebrate models. For example, they are readily available, as well as easy and inexpensive to maintain in the laboratory. Zebrafish can mate in minimal daylight, while many other fish require a

completely dark environment to lay eggs. They can also produce hundreds of offspring at weekly intervals and grow at an extremely fast rate, which means that they can be studied at any stage of their development. Moreover, zebrafish have a similar genetic structure to humans, with 70% of human genes found in their DNA; 84% of the genes have a human disease counterpart. In addition, being a vertebrate, the zebrafish has the same organs as humans, such as muscle, kidney and eye. Although the gene we pursue here is not present in humans, zebrafish is nevertheless a good model for studying basic biological processes common to many vertebrates and is an excellent model for the study of CSRP [69-74].

In this work, we specifically studied CSRP in zebrafish. The first goal of the study was to determine whether the recombinant zebrafish CSRP domains could function as Cu-binding SOD proteins to remove extracellular superoxide. We found that an isolated CSRP domain from zebrafish could bind Cu, and our preliminary data suggests that it has SOD activity. This is the first study to suggest SOD activity for CSRP of any organism.

The second goal for this project was to determine how CSRP is expressed in different zebrafish tissues and whether the CSRP protein is expressed in the fish as a full-length protein containing multiple tandem repeats of a SOD-like domain. We find that CSRP mRNA is most abundantly expressed in heart and tail and surprisingly, the protein appears to be expressed as 2x and 3x tandem repeats Cu-only SOD5 like domains rather than >100 kDa CSRP, suggesting that the full-length protein is processed into various units of Cu-only SOD domains. Interestingly, the zebrafish tissues that contain strong regeneration capacity due to ROS signaling are the heart and tail, which have the highest CSRP [61-64]. Therefore, the zebrafish is a good model for the study of CSRP and further studies are needed to see if CSRP has evolved in tissue regeneration, wound healing and

development.

2. MATERIALS AND METHODS

2.1 Expression of CSRP domains in *E. coli*

Plasmids PAG10H_ZfD1.1, PAG10H_ZfD1.2, PAG10H_ZfD3.1, PAG10H_ZfD3.2 and PAG_ZfD3.2, which were used to express CSRP domains either containing or lacking the 10X His tag, were kind gifts of Natalie Robinett. Recombinant CSRP proteins were produced in *E. coli* using plasmids encoding for zebrafish CSRP (residues 256-408 for ZfD1.1, residues 256-419 for ZfD1.2, residues 574-726 for ZfD3.1, and residues 574-734 for ZfD3.2) with the T7 promoter. His-tagged plasmids also contained an N-terminal 10X His tag and an intervening tobacco etch virus (TEV) protease cleavage site [53, 75].

In these studies, LEMO21 *E. coli* cells were used with pUC19 or CSRP plasmids for transformation of zebrafish CSRP domains. The tube containing LEMO21 *E. coli* cells was thawed on ice for 10 minutes. Next, 2 µl of control pUC19 or CSRP plasmid were added to the cell mixture and the tube was carefully flicked 4-5 times and placed on ice for 5 minutes. Following this step, 990 µl of room-temperature LB medium (5 g/L yeast extract, 10 g/L NaCl and 10 g/L tryptone) were pipetted into the mixture and the tube was placed at 37 °C and shaken vigorously at 250 rpm for one hour. After LB plates were warmed to 37 °C, 50-100 µl of each mixture was spread onto each plate and the plates were incubated overnight at 37 °C.

For the small-scale expression of CSRP domains, one colony from each plate was selected and added to 10 ml LB. This “starter” culture was shaken at 220 rpm at 37 °C overnight. Next, 10-20 µl of starter culture was added to another tube containing 10 or 50 ml LB and shaken at 220 rpm at 37 °C until the OD reached 0.6-0.8. At this density, 1 mM IPTG was added and the culture was shaken at 220 rpm at 37 °C for 4 more hours.

For the large scale expression, four separate flasks of 1 L each were prepared and the procedure for making the starter culture was repeated. Four flasks each of 1 L of TB buffer containing 24 g/L yeast extract, 12 g/L tryptone and 4 mg/L glycerol were prepared and autoclaved, and 17 ml of 1M KH_2PO_4 and 72 ml of 1M K_2HPO_4 were added right before expressing. Next, 10 ml start culture was added to each TB flask, together with 1 ml of 50 mg/ml carbenicillin into TB. The flask was shaken at 220 rpm at 37 °C until the OD reached 0.6-0.8, at which time the temperature was changed to 16 °C and the flask was cooled for 30 minutes. Following this, 1 mM IPTG was added and the flask was shaken at 200 rpm for 16-20 hours. Protein in all four flasks was later combined when harvesting the post-IPTG culture.

2.2 Purification and Cu-loading of CSRP domains in *E. coli*

The purification of CSRP domains used protocols adapted from Gleason et al and Peterson et al [53, 75] and the whole purification schemes are summarized in Fig. 4. After expression, the post-IPTG culture was centrifuged at 4000 rpm for 10 minutes. The pellet was saved and resuspended in ice-cold lysis buffer (50mM Tris, 200mM NaCl and 20mM Imidazole, pH 8.0) to obtain the total lysate. For the small-scale expression, 5 ml lysis buffer was added to resuspend the pellet. For the large-scale expression, 10 ml lysis buffer was added to resuspend the pellet from approximately 100 ml of post-IPTG culture. Next, inclusion bodies were made from the total lysate harvested. First, 1mg/ml lysozyme and 1mg/ml DNase were added to the resuspended total lysate, which was incubated for 30 minutes, then sonicated with Sonifier Cell Disrupters (VWR Scientific) to disrupt the cell membrane. During sonication, 4-8 cycles of 30 seconds ON/ OFF with 40% amplitude and 40-50% duty cycle were performed on ice, with the exact cycles depending on how heavy the duty was. After sonication, 2 µl of the culture was transferred to a microscope slide and covered with

microscope glass (ThermoFisher Scientific), then observed using the microscope. The cells were completely broken when the color of the mixture became clearer.

After sonication, the mixture was centrifuged at 25,000 x g for 30 minutes to isolate the inclusion bodies, which were resuspended in lysis buffer, centrifuged and washed twice with lysis buffer at 25,000 x g for 30 minutes, after which the pellet was collected. The pellet from around 100 ml of post-IPTG culture was resuspended in 100 ml of 50 mM Tris, pH 8.0, 8 M Urea with 1 mM DTT and stirred using a stir bar (Fisher Scientific) at 700 rpm at room temperature for 1-2 hours, after which time it was stirred at 4 °C overnight to unfold the protein. The unfolded protein was ultracentrifuged at 100,000 x g for 45 minutes and the supernatant, which contained unfolded protein, was saved and filtered using a 33 mm diameter sterile syringe filter with a 0.45-µm pore size hydrophilic Polyethersulfone membrane. The filtered solution containing refolded protein was dialyzed in 4 L of 50 mM Tris, pH 8.0 with 1 mM DTT for a minimum of 6 hours, after which time the dialysis solution was replaced and and the filtered solution was dialyzed for a minimum of 5 additional hours. The refolded protein was collected, ultracentrifuged at 100,000 x g for 30 minutes, and the supernatant containing crude refolded protein was saved.

Comparative experiments were performed following the same procedure, replacing 1 mM DTT with 1.5 mM reduced glutathione (GSH) during unfolding and replacing 1 mM DTT with 0.5 mM or 1 mM oxidized glutathione (GSSG) during refolding.

For the His-tagged protein, Ni column purification was performed. The 5 ml Ni column was first equilibrated with 50 mM Tris, pH 8.0, after which the protein was loaded and the column was washed with buffer A (50 mM Tris, pH 8.0). The protein was eluted with buffer B (50 mM Tris, 500 mM Imidazole, pH 8.0) at a flow rate of 1 ml/min. Fractions containing CSRP protein, as assessed

by SDS-PAGE, were combined and filtered through a 0.45-µm filter. Next, 1 mg TEV in 15 mg protein was added at room temperature and placed at 4 °C overnight to remove the 10X His tag. Protein with the His tag removed was loaded onto a 5 ml Ni column and the flow-through was collected at a flow rate of 4 ml/min, then loaded onto a Hi-trap Q column (GE Healthcare) also at the flow rate of 4 ml/min. The collected protein was assessed by SDS-PAGE. It should be noted that the non-His tagged protein did not need to undergo the Ni column purification.

Finally, Cu was loaded onto the purified protein in a process that was performed at 4 ^{o}C. First, the protein was dialyzed into the appropriate buffer. Four conditions were designed for Cu-loading. Conditions 1 and 2 were designed for protein refolded with 1 mM DTT and conditions 3 and 4 were designed for protein refolded with 1.5 mM GSSG. The quantity of each dialysis buffer was 4 L and 100 ml of protein were prepared in each condition

Condition 1: The protein was dialyzed in 50 mM Tris buffer, pH 8.0 for at least 8 hours to remove 1 mM DTT, then dialyzed overnight in 50 mM Tris, pH 8.0 with 0.25 mM Cu and finally dialyzed in 50 mM Tris buffer, pH 8.0 for at least 8 hours to remove excess Cu and remaining DTT.

Condition 2: The protein was dialyzed in 50 mM Tris, pH 8.0 with 1 mM DTT and 0.25 mM Cu overnight, then dialyzed at least 8 hours in 50 mM Tris, pH 8.0 with 0.25 mM Cu to remove DTT, and finally dialyzed in 50 mM Tris buffer, pH 8.0 for at least 8 hours to remove excess Cu and remaining DTT.

Condition 3: The protein was dialyzed for at least 8 hours in 50 mM Tris buffer, pH 8.0 to remove 0.5 mM GSSG, then dialyzed overnight in 50 mM Tris, pH 8.0 with 0.25 mM Cu, and finally dialyzed at least 8 hours in 50 mM Tris buffer, pH 8.0 to remove excess Cu and remaining GSSG.

Condition 4: The protein was dialyzed for at least 8 hours in 50 mM Tris buffer, pH 8.0 to

remove 1 mM GSSG, then dialyzed overnight in 50 mM Tris, pH 8.0 with 0.25 mM Cu, and finally dialyzed at least 8 hours in 50 mM Tris buffer, pH 8.0 to remove excess Cu and remaining GSSG.

For all four conditions, proteins were stored in the final buffer at 4 ^{o}C.

2.3 Biochemical Analysis of recombinant Zf CSRP domains

SDS-PAGE was carried out using Bolt 4-12% Bis-Tris Plus Gels (ThermoFisher Scientific) and 1X Bolt MES SDS Running Buffer (Novex). Precision Plus Protein All Blue Standards (Bio-Rad) of 10, 15, 20, 25, 37, 50, 75, 100, 150, and 250 kDa were used as references. All CSRP samples saved during expression and purification, such as total lysate, inclusion bodies, unfolded proteins, refolded proteins and Cu-loaded proteins, were added to 4X SDS sample buffer (50 mM Tris-HCl, pH 6.8, 8% SDS, 0.4% bromophenol blue, 40% glycerol) and heated at 100 °C for 10 minutes before loading. If a reducing gel was needed, 100mM DTT was also added before loading. Gels were stained by Coomassie Brilliant Blue for 30-60 minutes and then destained in 40% methanol and 10% acetic acid in Milli Q water. The destained gels were visualized using an Odyssey imaging system (LI-COR Biosciences).

For Cu analysis, protein treated with Cu and dialyzed following each of the above four conditions was analyzed for Cu content using atomic absorption spectroscopy (AAS) performed on an AAnalyst600 graphite furnace atomic absorption spectrometer (PerkinElmer Life Sciences). The total protein concentration was determined using a nanodrop, and the Cu content was compared with the protein concentration to obtain the Cu/total protein ratio. The fraction of total protein represented by CSRP (≈70%) was determined by quantifying Coomassie stained bands through image studio software.

For SOD activity analysis, 10 or 40 µg of the CSRP domain-containing sample was subjected to

native gel electrophoresis at 50 mA using 10% Tris–glycine native gels (Novex). SOD activity was monitored by nitro blue tetrazolium (NBT) staining and Coomassie Brilliant Blue staining [75, 76].

To compare anti-SOD5 versus anti-ZfD1D3 reactivity with recombinant CSRP domains, 1 μg of each sample was subjected to denaturing gel electrophoresis on a 4-12% Bis-Tris Plus Gel (ThermoFisher Scientific), followed by immunoblotting using anti-SOD5 antibody (1:1000) or anti-ZfD1D3 (1:500), and then a secondary anti-rabbit antibody (1:10,000, Alexa Fluor 680). CSRP immunoreactivity was visualized using an Odyssey imaging system (LI-COR Biosciences). The anti-SOD5 antibody from anti-SOD5 rabbit serum (JH929) was a gift from Ryan L. Peterson [75]. The anti-ZfD1D3 rabbit serum was created using purified recombinant *E. coli* ZfD1.2 andZfD3.2 proteins as described above and a 90-day rabbit protocol (Cocalico Biological, Inc), and the antibody we used was test bleed 2 antibody.

2.4 Extraction of CSRP protein and RNA from Zf tissues

Zebrafish tissues chosen in this experiment included gills, brains, hearts, kidneys, intestines and tails from both male and female adult (6-month-old) fish, as well as male testes and female ovaries. 28 hpf embryos were also used for study. CSRP protein and RNA were both isolated from zebrafish tissue using the TRIzol protocol. First, 1ml TRIzol reagent was added to each tube. For the protein extraction protocol, the same tissue from 3 fish of the same gender were collected and placed into one tube, while for the RNA extraction protocol, only one tissue from 1 fish was placed into one tube. Alternatively, 50 of the 28 hpf embryos were collected and placed into one tube. The tissues were homogenized in each tube using a homogenizer (Kimble, Pellet Pestles 749540-0000 Drive Unit Cordless Motor) which uses a battery-operated micro DC motor and has a fixed speed of 13,000 rpm. The homogenized sample was incubated for 5 minutes at room temperature

to allow complete dissociation of tissue material. Next, 0.2 ml chloroform per 1 ml TRIzol Reagent used for homogenization was added and the tube was capped securely and shaken vigorously by hand for 15 seconds. The sample was incubated for 2-3 minutes at room temperature and then centrifuged at 12,000 x g for 14 minutes at 4 °C. Following centrifugation, the mixture was separated into a red lower phenol-chloroform phase, an interphase, and a colorless upper aqueous phase. The RNA remained exclusively in the aqueous phase, while the protein remained in the lower red phenol-chloroform phase. The two phases were separately saved in new tubes.

For the protein isolation protocol, 1.5 ml of isopropanol was added to each tube containing the phenol-chloroform phase. The sample was incubated for 10 minutes and centrifuged for 10 minutes at 12,000 x g at 4 °C. The protein remained in the pellet, and the supernatant was discarded using a micropipettor. Next, a wash solution consisting of 0.3 M guanidine hydrochloride in 95% ethanol was prepared, and the pellet was resuspended in 2 ml of this wash solution and incubated for 20 minutes. The mixture was then centrifuged for 5 minutes at 7500 x g at 4 °C and the pellet was saved. This resuspending-centrifuging step was repeated twice, after which 2 ml of ice-cold 100% ethanol was added and mixed by vortexing briefly. The mixture was incubated for 20 minutes at room temperature and centrifuged for 5 minutes at 7500 x g at 4 °C, after which the protein remained in the pellet. The supernatant was discarded and the protein pellet was air dried for 5-10 minutes. The pellet was resuspended in 200 µL of 10 M urea by pipetting up and down, and the sample was incubated at 50 °C in a water bath or heat block to ensure complete resuspension of the pellet. Following this, the sample was centrifuged for 10 minutes at 10,000 x g at 4 °C to remove insoluble materials. The protein was in the supernatant, which was transferred to a new tube and stored at -20 °C for later analysis.

For the RNA isolation protocol, 0.5 ml of 100% isopropanol was added to the aqueous phase for homogenization. The sample was incubated at room temperature for 10 minutes, then centrifuged at 12,000 x g for 10 minutes at 4 °C and the supernatant discarded. The pellet was washed with 1 ml of 75% ethanol and vortexed briefly, then centrifuged at 7,500 x g for 5 minutes at 4 °C and the wash discarded. The pellet was air dried for 5-10 minutes and resuspended in RNase-free water (30 uL) by passing the solution up and down several times through a pipette tip. The solution was incubated in a heat block at 55-60 °C for 10-15 minutes to completely dissolve RNA and then stored at -80 °C for later analysis.

2.5 Biochemical Analysis of CSRP protein and RNA from Zf tissues

For RNA analysis, zebrafish tissue RNA was isolated using the TRIzol protocol as described above. The RNA concentration was determined using a BioTek Synergy HT plate reader. Next, RNase-free water was added to the RNA samples to 100 µl total and the RNA samples were treated with DNase using the RapidOut DNA Removal Kit (ThermoFisher Scientific). cDNA was prepared from 1 µg of RNA using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific). cDNA was subsequently diluted to 1:5 and analyzed by qRT-PCR using 2X PowerUp SYBR Green Master Mix (ThermoFisher Scientific). The CSRP expression levels were calculated using the ΔCT method (where CT is threshold cycle). A lower ΔCT value indicates a higher gene expression. Expression levels are shown in comparison to the housekeeping gene *actb1* set at the value of 1. The two *actb1* primers are kind gifts of Ed Culbertson, who also helped design the CSRP primers using SNAPGENE. Primers for CSRP are located in Exon 4 and 5 of the gene. Primers used are described in table:

Gene name		Primer Sequence

actb1	Forward	CGA GCA GGA GAT GGG AAC C
	Reverse	CAA CGG AAA CGC TCA TTG C
CSRP	Forward	ATG GCA GTT ACG GTG ATG TT
	Reverse	CCA TCT GGG TAC ATG GGT TTA

To determine how CSRP was expressed in different zebrafish tissues and embryos, protein concentration was tested using a nanodrop. Next, a certain amount of each sample was subjected to denaturing gel electrophoresis on a 4-12% Bis-Tris Plus Gel (ThermoFisher Scientific), followed by immunoblotting using anti-SOD5 antibody (1:1000) , anti-ZfD1D3 (1:500), or control anti-beta actin (1:5000), and then a secondary anti-rabbit antibody (1:10,000, Alexa Fluor 680). CSRP immunoreactivity was visualized using an Odyssey imaging system (LI-COR Biosciences).

3. RESULTS AND DISCUSSION

PART 1: The expression and biochemical analysis of recombinant *D. rerio* CSRP domains

3.1.a Both domain 1 and domain 3 of the zebrafish *Danio rerio* Cu-only SOD repeat protein can be expressed in *E. coli*.

As mentioned above, zebrafish CSRP has three perfect domains with a Cu-binding site. In this thesis, we chose to focus on domains 1 and 3 because each only has a single pair of cysteines forming the disulfide bond, while domain 2 has multiple cysteines which might cause problems during purification by forming disulfide crosslinks. We made two constructs for each domain of different sequence lengths with the longer construct containing approximately 10 additional C-terminal amino acids, which may help with proper folding, protein stability and activity. These four domains were named ZfD1.1, ZfD1.2, ZfD3.1 and ZfD3.2.

Based on this, it was first sought to determine whether the four individual constructs could be expressed in *E. coli*. The four individual domains with the 10x His tag were tested in an expression plasmid which was induced by IPTG. The result shows that all four domains can be expressed in *E. coli*, as seen in the Coomassie staining of the total lysate (Fig. 5A). Subsequent studies on the His-tagged proteins focused on ZfD3.1, based on its high-level expression (Fig. 5A).

Previously published studies on Cu-only SOD5 showed that the protein expressed in *E. coli* accumulated in inclusion bodies [53, 75]. We next determined whether this is also true for our Zf CSRP domains. We performed another Coomassie staining of both ZfD3.1 total lysate and inclusion bodies (Fig. 5B). The gel shows that in inclusion bodies, there is a clear band at around 18.7 kDa, which confirmed that ZfD3.1 also accumulates in inclusion bodies. In addition, two bands of ≈38 and 14 kDa are often seen, reflecting non-specific proteins from *E. coli* inclusion bodies (see Fig. 11).

3.1.b All four His-tagged zebrafish CSRP domains form oligomers upon refolding.

The aforementioned study indicates the accumulation of ZfD3.1 in inclusion bodies. Next, we sought to further purify the ZfD3.1 inclusion bodies. We used the protocol in which the protein was unfolded and refolded in the presence of DTT and subject to Ni affinity chromatography (Fig. 4). It was found that ZfD3.1 formed extensive oligomers during the refolding process, as seen under nonreducing conditions (Fig. 6A, lanes 3-4 and Fig. 6B, lane 3). The molecular weight of these oligomers correspond to multimers of the ≈19 kDa ZfD3.1 domain of 38 kDA (dimer), 57 kDa

(trimer), 76 kDa (tetramer) etc. Reducing conditions did not completely resolve these oligomers (Fig. 6A, lane 2) unless the samples were boiled (Fig. 6B, lane 4). This indicates that ZfD3.1 forms intermolecular disulfide crosslinks during refolding that in part contribute to oligomerization. Interestingly, such oligomerization is not typical of *C. albicans* SOD5, which forms a monomer under all conditions [53].

When applied to the Ni column, virtually all of the protein eluted in the flow-through fraction and none was seen to adhere to the column (Fig. 6B, lane 7-12). We also tested TEV cleavage of all fractions that eluted from the Ni column, which failed to work as well (data not shown). Overall, these results suggest that the oligomers may be aggregating, preventing adherence to the Ni column.

In addition, we tested whether the other three His-tagged domains form oligomers as does ZfD3.1, using the same purification scheme as described for ZfD3.1 in Figs. 5-6. An additional filtering step was performed to aid in the dissociation of the oligomers. The results indicate that all three domains accumulated in inclusion bodies and formed oligomers upon refolding (Fig. 7A-C). Most of the ZfD1.1 and ZfD1.2 was lost after filtering, perhaps due to aggregation or adhering to the dialysis membrane, while ZfD3.2 still seemed to aggregate after filtering (Fig. 7A-C). Overall, these results indicate that for all four His-tagged constructs, the protein forms large oligomers or aggregates.

The protocol was then modified. Considering that inclusion bodies already represented a substantial purification without Ni column purification (Fig. 7A-C, lane 3) combined with the possible contributions of adding a His tag to misfolding and aggregation, we decided to design new zebrafish CSRP domains without the 10X His Tag.

3.1.c ZfD3.2 without His tag is less prone to aggregation during refolding and shows activity in the native gel assay when reconstituted with Cu in the presence of GSH.

A construct for expressing ZfD3.2 without the His tag was prepared. The same steps for making total lysate and inclusion bodies of new ZfD3.2 were followed. We found out that the new ZfD3.2 protein still accumulated in inclusion bodies (data not shown). Previously published studies on Cu-only SOD5 showed that proper refolding and Cu-loading of SOD5 could be achievevd in the presence of glutathione [53, 75]. Therefore, we tried two methods for refolding the protein, as is shown in Fig. 4. In the first method, the protein was unfolded and refolded in the presence of DTT,

as was done for the purification of His tagged Zf CSRP. In the second method, DTT was replaced with GSH (reduced) in the unfolding step and GSSG (oxidized) in the refolding step. The entire process was conducted at pH 8.0 because ZfD3.2 is not stable and can precipitate at a pH lower than this (data not shown). It was observed that ZfD3.2 without the His tag is less prone to aggregation during refolding in the presence of DTT compared to 10X His tagged ZfD3.2, based on the abundance of monomeric versus oligomeric protein (Fig. 7C, lane 5 for His-tagged and Fig. 8, lane 9 for no His tag). Additionally, more oligomers appear to form with the untagged protein in the presence of DTT compared to GSH/GSSG (Fig. 8, lanes 7-9 compared to lanes 2-6). However, the intramolecular disulfide of ZfD3.2 cannot be completely oxidized during refolding, even with a higher concentration of GSSG (Fig. 8, lanes 3-4).

We attempted to load the CSRP domain with Cu by dialyzing against Cu in the presence of DTT or oxidized GSSG, followed by extensive dialysis to remove unbound Cu. This treatment resulted in some increase in oligomer formation; however, the oligomerization was greater in the presence of DTT (Fig. 8, lanes 7 and 8) compared to GSSG (Fig. 8, lanes 5 and 6). We then used Cu atomic absorption spectroscopy to determine whether the sample retained Cu following dialysis. The results in Table 1 show that the ratio of Cu retained to total protein is over 0.7 in the presence of GSSG, which is much more than the ratio obtained in the presence of DTT. Based on our estimation using imaging software to quantitate bands, our protein samples are roughly 70% Zfd3.2 (Fig. 8, lane 1). Our findings are consistent with a 1:1 ratio of Cu to protein with ZfD3.2, although this awaits further purification of the protein.

Next, we examined whether our Cu-loaded sample containing ZfD3.2 showed SOD activity. We used a native activity gel based on superoxide reduction of NBT which had previously been used for SOD activity including Cu-only SODs [76]. The sample containing ZfD3.2 was run in native gels with SOD5 as a control. Following this, one gel was stained with NBT stain and the other with Coomassie. The result shows that the protein prepared with DTT has no detectable activity (Fig. 9A, top gel). One explanation for this is that the protein did not enter the native gel, as can be seen from the Coomassie-stained gel (Fig. 9A, bottom gel). However, it was found that the protein prepared with GSH/GSSG did exhibit some SOD activity (Fig. 9B, top gel). This protein was able to enter the Coomassie-stained native gel (Fig. 9B, bottom gel) and this correlated with SOD activity as measured by SOD staining. Although it appeared that the activity from the CSRP-containing

sample is roughly 4-fold less than that of Cu-only SOD5, it has SOD activity nevertheless. This is a very exciting preliminary result because it suggests that a CSRP domain has SOD activity.

3.1.d Points for Discussion

In this study, we found that unlike *C. albicans* Cu-only SOD5 [53], Zf CSRP tends to form oligomers. The reason for this most likely has to do with the way CSRP is naturally present as a four-tandem-domain repeat. Therefore, it is possible that the individual domains of CSRP interact with one another while *C. albicans* SOD5 tends to be monomeric. In the future, we can attempt to express and purify the full-length protein and analyze by structural biology whether the four domains exist as interacting repeats or the polypeptide is in more of an extended conformation (Fig. 14). In addition, His-tagging the CSRP seems to increase the presence of oligomers, and the reasons for this are unknown. It may be that the His tag itself affects the stability of the protein and increases protein-protein interactions.

We also found that Zf CSRP is less prone to aggregation when refolded in GSSG than in DTT. This result may be due to the fact that GSSG is an oxidizing agent and can promote intramolecular disulfides in the protein, while DTT is a reducing agent and does not allow CSRP to be effectively oxidized. This may cause some of the single domains to end up in disulfide crosslinks and form more oligomers. In addition, since the disulfide is a necessary for Cu-loading and protein function [32, 53], the inability to form a disulfide bridge in the DTT-prepared protein could explain why the protein binds Cu poorly and is not active.

This study is the first to suggest that the Zf CSRP domain can bind Cu and has the potential to act as a SOD (Fig. 14). Future studies can possibly examine whether the full-length recombinant CRSP protein can act as a SOD, or if GPI-anchored CSRP in live cells can also function as SOD protein to remove extracellular superoxide produced through NOX enzyme. The main limitation of this study is that all of the experiments were conducted with impure inclusion bodies which contained other proteins. Therefore, additional purification, such as ion exchange and HPLC as has previously been done for Cu/Zn SOD proteins [77-79], will be required in future studies.

PART 2: Tissue specific expression of CSRP in adult zebrafish and embryos

3.2.a CSRP mRNA expression levels in zebrafish vary in different tissues and embryos.

Based on the interesting results from the recombinant protein experiments, we then wished

to examine the expression of CSRP in zebrafish. First, we wished to ascertain the CSRP mRNA levels in zebrafish embryos and 6-month-old adult tissues. The tissues which we selected for study included gill, brain, heart, kidney, intestine, tail, ovary and teste. The fish were dissected and male and female tissues were separated for analysis by qRT-PCR. In Fig. 10, expression levels are shown in comparison to the housekeeping gene *actb1* and each bar represents an individual fish tissue. From Fig. 10A, it can be seen that in male fish, the expression level is highest in the heart and the second highest is in the tail, followed by the levels in the gill and brain, while the level in the intestine is virtually undetectable. As seen in Fig. 10B, the expression level in female fish is highest in the heart, followed by the level in the tail and gill. The expression level in the brain is also high, while the levels in the kidney, ovary and intestine are barely detectable. In addition, we can see that in both male and female fish, the levels of expression are similar; the highest level of expression is in the heart, followed by the levels in the tail, gill and brain, while the levels in the intestines are virtually undetectable.

3.2.b Validating antibodies for the study of zebrafish CSRP in tissues by immunoblot

Since the RNA analysis was not definitive proof of tissue expression, we decided to further look at the protein by immunoblot. Because there was no antibody commercially available or published for CSRP, we had to formulate our own. As introduced in Fig. 3, CSRP is a SOD5 like protein, so two approaches were attempted: the first using previously published anti-SOD5 antibody and the second using one that we made ourselves against ZfCSR domains 1 and domains 3, collectively named anti ZfD1D3 antibody. We ran SDS reducing gels of ZfD3.1 and ZfD3.2, and then determined by western blotting if ZfD3 can be recognized by the antibodies. We also subjected the gel to Coomassie staining before western blotting (Fig. 11A) to compare immunoreactive protein to total ZfD3.1 and ZfD3.2 protein. From the western blots, we can see the specific ZfD3.1 band at around 18 kDa and the specific ZfD3.2 band at around 20 kDa in both gels, which indicates that ZfD3 can be recognized by both anti-SOD5 (Fig. 11B) and anti-ZfD1D3 (Fig. 11C). However, comparing the same bands in the two western blots, the signal-to-noise ratio was much better with anti-SOD5. Despite this, we chose to use both antibodies in our studies of CSRP protein expression in zebrafish tissues.

3.2.c The analysis of CSRP protein expression in zebrafish embryos and adult tissues

Next, we wished to test the CSRP protein expression levels in zebrafish embryos and adult

tissues. Forty embryos at 28 hpf (hours post fertilization) were combined, as were the same tissues from three male and female fish. The tissue and embryos were then analyzed by western blotting using either anti-SOD5 antibody or anti-ZfD1D3 antibody. Results from analysis of the male tissues are shown in Fig. 12. Western blotting with anti-SOD5 antibody shows that the male heart, which has the highest level of CSRP expression based on the qRT-PCR experiment (Fig. 10A), contains an abundant cross-reactive species of 40 kDa that is unique to heart and not present in any other male tissue (Fig. 12A, top). The abundance of this 40-kDa species in the heart may be underrepresented based on the apparent loss of sample, indicated by beta actin (Fig. 12A, bottom). Interestingly, this 40-kDa species cannot be detected by anti-ZfD1D3 antibody (Fig. 12B). The reason for this is unknown, but it may be that our anti-ZfD1D3 antibody does not cross-react well with purified CRSP domains compared to anti-SOD5 antibody (Fig. 11C). The tail expresses the second-highest level of CSRP as shown by qRT-PCR (Fig. 10A), and western blotting with anti-SOD5 shows a doublet at >100 kDa and a minor species at around 60 kDa (Fig. 12A). The 60-kDa band also appears in tail tissue when anti-ZfD1D3 is used (Fig. 12B). Similar results were observed in females, with a major 40-kDa species in the heart seen when anti-SOD5 is used and ≥100-kDa and 60-kDa species in the tail seen when both anti-SOD5 and anti-ZfD1D3 are used (Fig. 13). Additional species appeared in other female tissues. For example, from qRT-PCR analysis, it can be seen that female brain tissue shows an intermediate level of CSRP mRNA expression (Fig. 10B); from western blotting, we can also see a 100-kDa species in brain with anti-SOD5 and a 60-kDa species in brain with anti-ZfD1D3. The qRT-PCR analysis shows virtually undetectable CSRP expression in the ovary and female intestine, and the 60-kDa band is shown to be absent in both of these tissues when anti-SOD5 antibody and anti-ZfD1D3 antibody are used. Instead, ovary and intestine tissues contain 50-kDA and 18-kDA reactive species which cross-react with anti-SOD5 and are of unknown nature. It is important to note that 40 and 60 kDa bands specific to heart and tail correspond to the molecular weight of 2x and 3x tandem repeats of SOD5-like domains, suggesting that the protein is processed into smaller units of Cu-only SOD domains.

3.2.d Points for Discussion

From the mRNA expression of CSRP in adult zebrafish and embryos, we know that, in both males and females, CSRP is most abundantly expressed in the adult heart and to a lesser degree in the tail. CSRP mRNA is essentially absent from other tissues, including intestine, kidney and the

reproductive organs. Interestingly, the same two tissues in fish with high CSRP are those with strong regenerative capacity [61-64], an event closely tied to ROS. The ability to regenerate tissue greatly varies among animals, and in those animals which retain regenerative capacity as adults (marine invertebrates, insects, teleosts and amphibians), tissue regeneration has been associated with large bursts of ROS [80-83]. The NOX enzymes and partner Rho GTPases which are responsible for this ROS have been identified [81, 82, 84, 85], and the downstream targets of the ROS which mediate tissue regeneration have been characterized [82, 86, 87], but nothing is known about the control of NOX superoxide. The ROS signal for tissue regeneration is H_2O_2, not superoxide, and conversion of NOX superoxide to H_2O_2 requires a SOD. In mammals, NOX enzymes partner with an extracellular Cu/Zn SOD tethered to the cell surface through GPI anchors [88], and while zebrafish do express an extracellular Cu/Zn SOD, it lacks the sequences for cell surface attachment and is completely absent in the heart. We hypothesize that in the fish heart, and perhaps other tissues including the fin and tail, Zf CSRP acts as a SOD to control NOX superoxide for tissue development, repair and regeneration. If this is true, CSRP will be the first protein to see superoxide emanating from NOX during tissue regeneration. In our future studies, we can test the importance of CSRP in tissue regeneration of the tail and fin using zebrafish models in collaboration with Professor Andrew McCallion at the Johns Hopkins School of Medicine zebrafish facility.

From the study of CSRP protein expression by immunoblot, it was learned that CSRP is most abundant in zebrafish heart and tail tissues. However, the bands we observed were 40 and 60 kDa, which do not correspond to the correct molecular weight for full-length CSRP (around 100 kDa). The reasons for not detecting full length protein are unknown and, as one possibility, protein glycosylation or aggregation of the full-length protein may prevent the recognition of CSRP by the antibodies used or the entry of the protein into the gel. In these experiments, the protein was not deglycosylated and in future experiments, the effects of deglycosylation and perhaps various detergents to release GPI-anchored proteins can be tested. Since only small proteins can be seen with anti-SOD and anti-CSRP, these 40- and 60-kDa species may represent clipping of the intact polypeptide, as depicted in Fig. 14. Interesting, a similar clipping has been reported for the intestinal lactase of mammals. Like CSRP, lactase appears to have evolved from twice duplication of a lactase enzyme, resulting in four tandem repeats of the enzyme on a single polypeptide. The mature enzyme however, only has one pair of tandem repeats, the result of controlled processing

in the secretory pathway [89]. A similar situation may be true for CSRP.

Finally, it needs to be verified that the bands seen on the western blots are actual CSRP. We know that our antibodies do not cross-react with Cu/Zn SOD1 (Sabrina Schatzman, personal communications), but we cannot exclude the possibility that there might be other cross-reactivity between the antibodies used and non-CSRP zebrafish proteins. In future studies, a knock out or knock down of CSRP will be created in order to verify CSRP recognition by our antibodies.

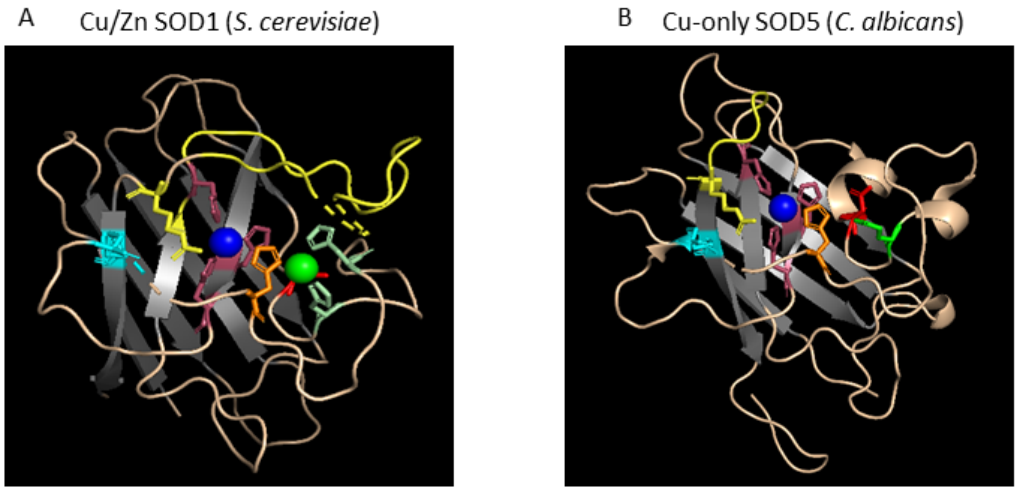

Fig. 1 The active site of Cu/Zn versus Cu-only SODs.
Three-dimensional structure of *Saccharomyces cerevisiae* Cu/Zn-SOD1 (A) and *Candida albicans* Cu-only SOD5 (B) as determined by PyMOL. Key features are labeled with different colors: Backbone sequences are in wheat, the Greek key β barrel is in grey, the electrostatic loop and active site arginine are in yellow, disulfide cysteines are represented as cyan sticks, Cu as a blue sphere and Zn as a dark green sphere, the dynamic histidine as an orange stick, the other Cu binding histidines as pink sticks, the Zn binding histidines as light green sticks, SOD5 Glu-110 as a green stick and Asp-113 as a red stick.

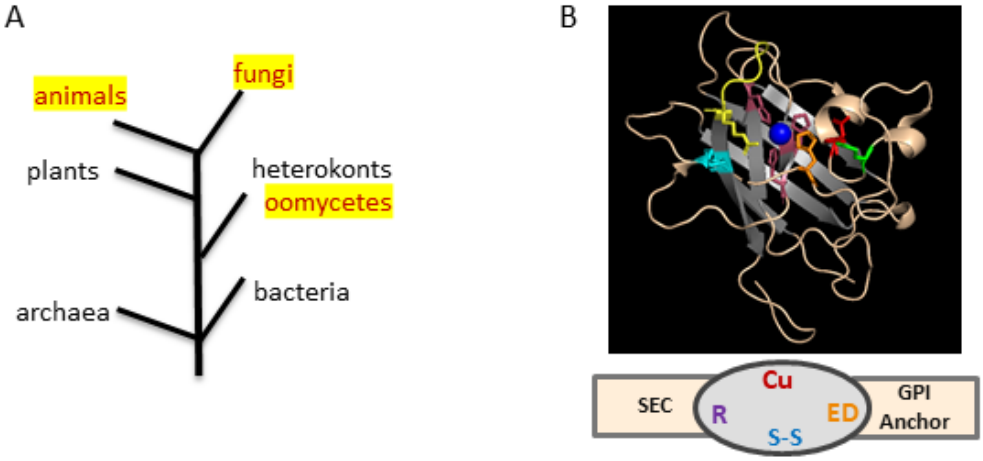

Fig. 2 Evolution of Cu-only SOD Domains.

(A) Existence of Cu-only SOD5 like domains (red and yellow highlights) shown in the phylogenetic tree. (B) Top,three-dimensional structure of *C. albicans* SOD5 as determined by PyMOL; bottom, schematic of the predicted full length of *C. albicans* SOD5.N and C terminal sequences and GPI anchorage are in wheat; β-barrel domains are ovals with active site arginine (R), Cu site (Cu), disulfide cysteines (S-S) and active site Glu-110 and Asp-113 (ED).

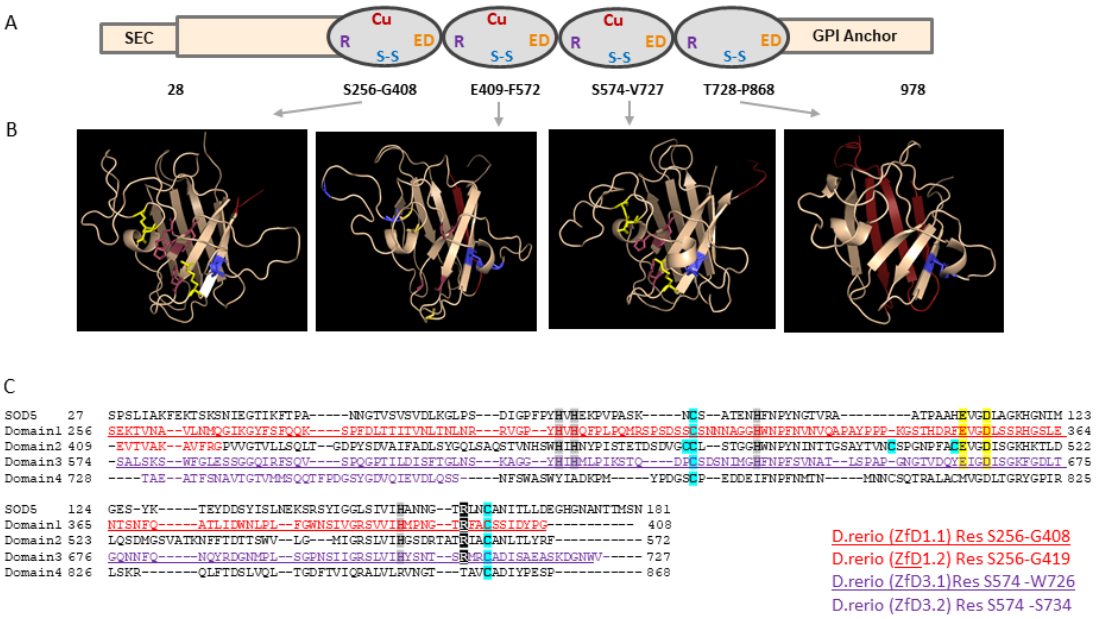

Fig. 3 The four SOD5-like Cu-only SOD repeat protein (CSRP) domains of the zebrafish *Danio rerio*. (A) Schematic of the predicted full length of zebrafish *Danio rerio* Cu-only SOD repeat protein; the N and C terminal sequences for secretion and GPI anchorage are in wheat, the Greek key β-barrel domains are shown in ovals with active site arginine (R), Cu site, disulfide cysteines (S-S) and sequences corresponding to the active site Glu-110 and Asp-113 of *C. albicans* SOD5 (ED). (B) Predicted three-dimensional structures of the four SOD5-like Cu-only SOD repeat protein domains as determined by PyMOL; the most C-terminal domain (ZfD4) lacks a Cu site but retains the predicted overall fold as well as other necessary features of Cu-only SOD5-like domains. The overlapping sequences of our designed ZfD1.2 with Zf domain2 and the overlapping sequences of our designed ZfD3.2 with Zf domain4 are shown in red. (C) Alignment of the four individual SOD5-like CSRP domains against *C. albicans* SOD5; the Cu binding histidines are in gray, disulfide in blue, active site arginine in black and positions equivalent to SOD5 Glu-110 and Asp-113 in yellow. The overall amino acid identity and similarity compared to SOD5 is as follows: ZfD1: 33% identity and 51% similarity, ZfD2: 30% and 47%, ZfD3: 31% and 56%, ZfD4: 27% and 43%. The amino acid sequences of the two constructs of Zf domain1 which we designed (ZfD1.1 and ZfD1.2) are in red; the amino acid sequences of the two constructs of Zf domain3 which we designed (ZfD3.1 and ZfD3.2) are in purple.

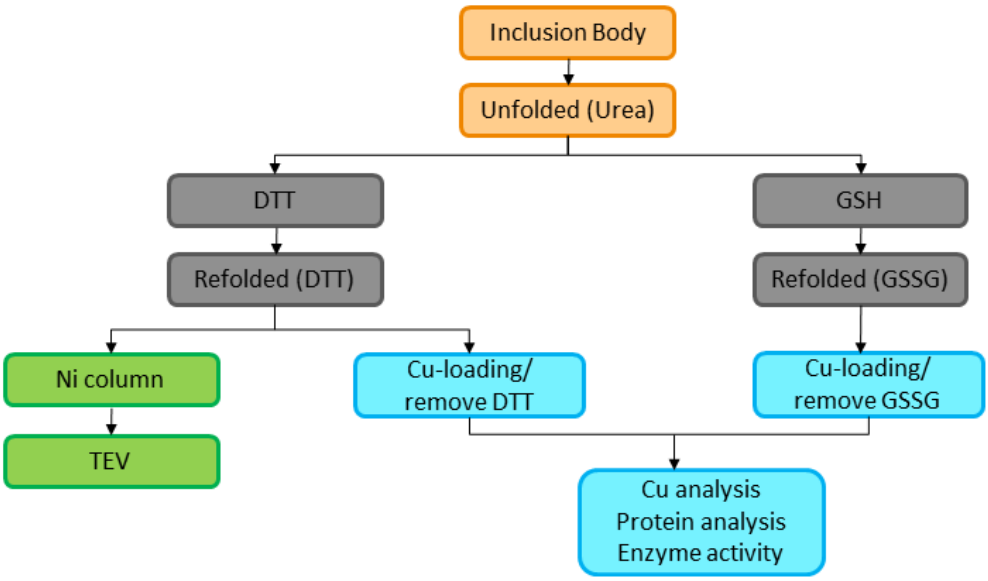

Fig. 4 Purification and Cu-loading scheme for Zf CSRP domains.

CSRP expressed in E. *coli* were accumulated in inclusion bodies. The first purification step was the unfolding of the protein inclusion bodies in urea, as marked in orange. The grey rectangles are the two methods used for protein unfolding and refolding: Either DTT in urea was added and the protein was refolded in the presence of DTT, or GSH in urea was added, followed by the addition of GSSG to refold CSRP. The His-tagged Zf CSRP samples were further purified using a Ni column and TEV cleavage, as marked in green. For the Zf CSRP without His tag, we directly loaded Cu and performed further Cu and protein analysis, as shown in blue.

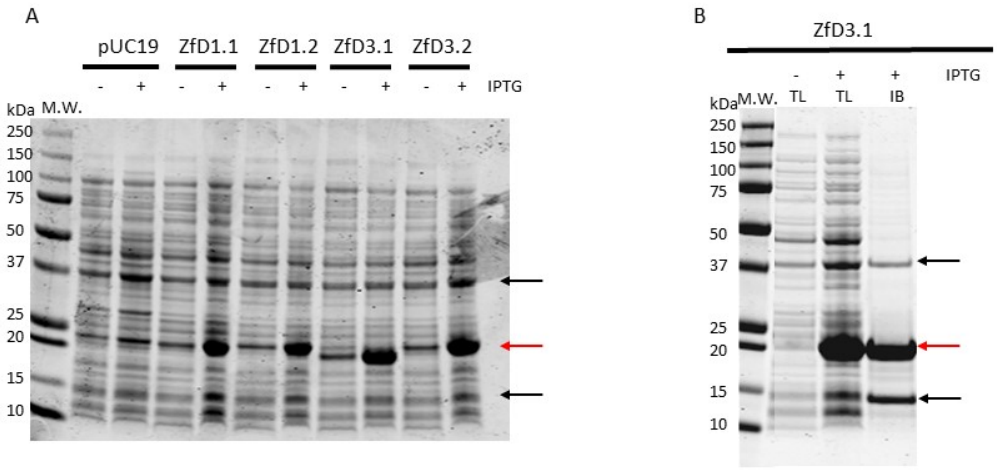

Fig. 5 Expression of ZfD1.1, ZfD1.2, ZfD3.1, ZfD3.2 in *E. coli* as monitored in whole cell lysates and the presence of ZfD3.1 in inclusion bodies.

(A) 10 µg of whole cell *E. coli* protein from cells transformed with either the pUC19 control or the indicated domains of Zf CSRP were run on a SDS reducing gel and stained with Coomassie. The red arrow indicates the position of Zf CSRP. The molecular weight of ZfD1.1 is ~19.7 kDa, that of ZfD1.2 is ~20.8 kDa, that of ZfD3.1 is ~18.7 kDa, and that of ZfD3.2 is ~20.2 kDa. (B) 10 µg of total cell lysate (TL) and inclusion body (IB) *E. coli* protein from cells transformed with ZfD3.1 were run on a SDS reducing gel and stained with Coomassie. The red arrow indicates the position of ZfD3.1 with a molecular weight of ~18.7kDa and the two black arrows indicate the non-specific bands with molecular weights of 38 kDa and 14 kDa. M.W., molecular weight.

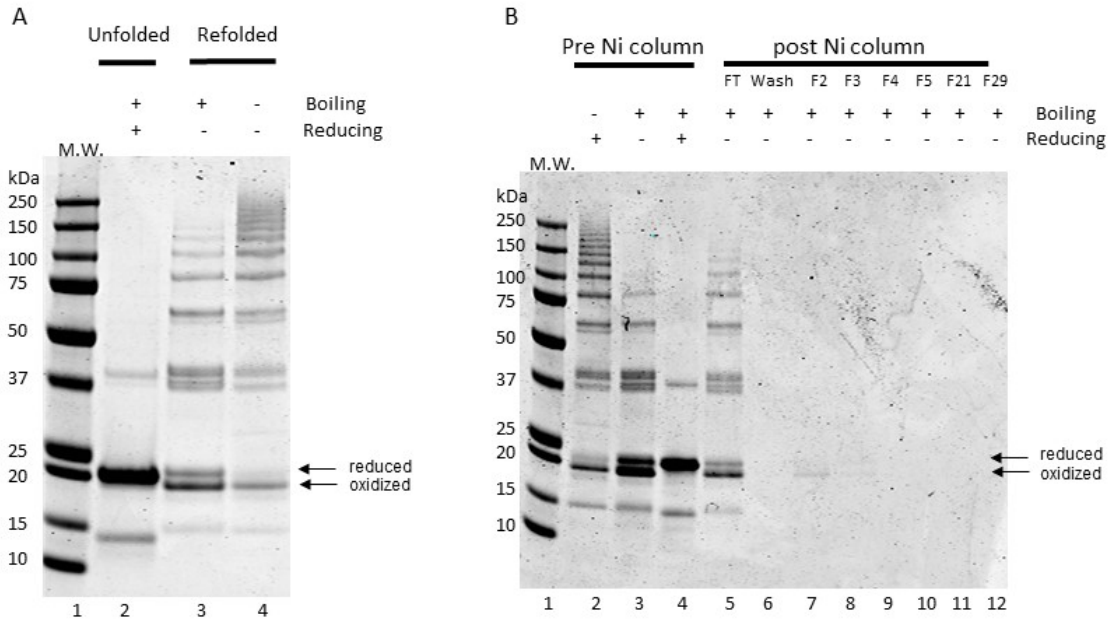

Fig. 6 Purification of His-tagged ZfD3.1.

(A) Unfolded and refolded ZfD3.1 in SDS gel with Coomassie stain. 20 µl of each sample or identical cell equivalents were loaded onto the gel (see Materials and Methods). Boiling (+) indicates that the samples were heated at 100°C ten minutes before the gel was run; reducing (+) indicates that extra DTT was added to the samples before the gel was run. Arrows indicate positions of reduced or oxidized ZfD3.1. (B) SDS gel of refolded ZfD3.1 and fractions collected post-Ni column purification with Coomassie stain. 20 µl of each sample or fraction were loaded onto the gel. FT, flow-through; Wash, column washing after flow-through; F2/3/4/5/21/29, number fractions which eluted from the column by adding buffer B (50mM Tris/500mM Imidazole, pH 8.0).

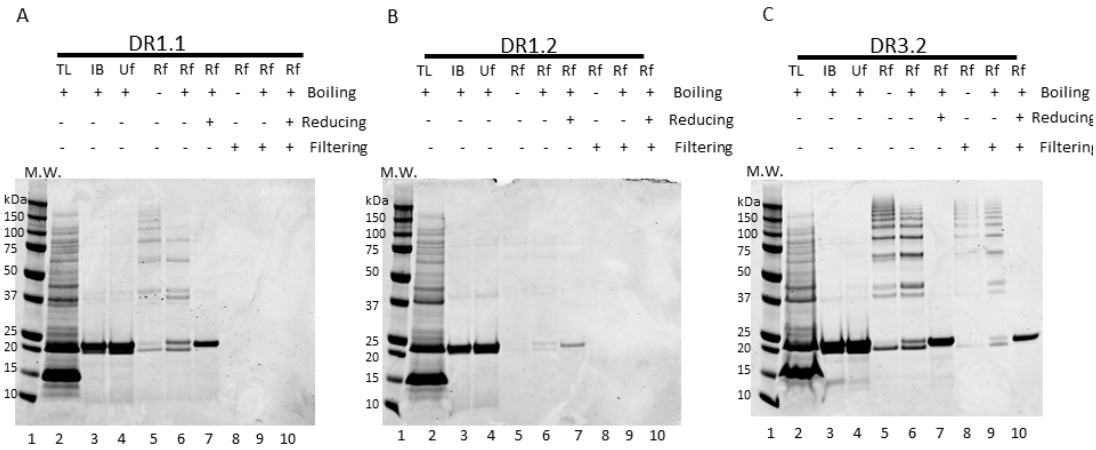

Fig. 7 Expression and purification of ZfD1.1, ZfD1.2 and ZfD3.2.

Coomassie-stained SDS gels of samples obtained from *E. coli* expressing the indicated CSRP isolated domains. TL, total lysates; IB, inclusion bodies; Uf, unfolded proteins; Rf, refolded proteins. Boiling, samples were heated to 100 ^{o}C 10 min before gel was run; reducing, extra DTT was added to samples before gel was run; filtering, further purification of refolded samples using 33 mm diameter sterile filter was performed. M.W., molecular weight. The molecular weight of ZfD1.1 is ~19.7kDa, that of ZfD1.2 is ~20.8kDa, and that of ZfD3.2 is ~20.2kDa. Samples represent whole cell and inclusion body *E. coli* protein from cells transformed with ZfD3.2 and the resulting purified unfolded and refolded proteins. +, sample undergoes treatment; -, sample does not undergo treatment. The same amount of each sample was loaded (15 µl; see Materials and Methods).

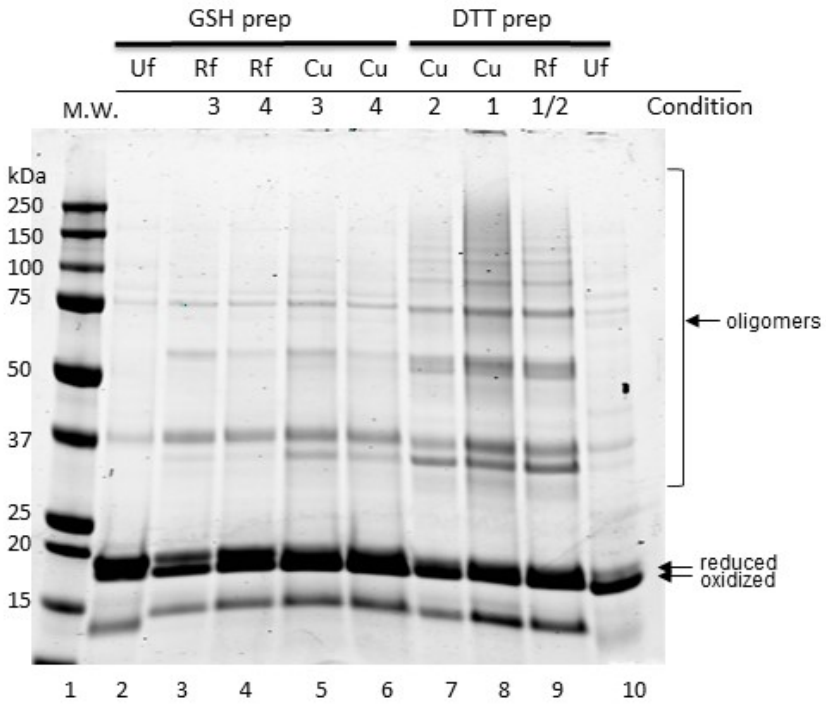

Fig. 8 Expression, purification and Cu-loading of ZfD3.2 without His tag.

Two methods were used for unfolding, refolding and Cu-loading of ZfD3.2: the first using DTT ("DTT") and the second with glutathione added during the process ("GSH"). See Materials and Methods for details. Samples were run on a SDS non-reducing gel, which was stained with Coomassie. Uf, unfolded; Rf, refolded; Cu, Cu-loaded D3.2; Condition 1 indicates the removal of 1 mM DTT followed by the addition of Cu; Condition 2 indicates the addition of Cu followed by the removal of 1 mM DTT; Condition 3 indicates the removal of 0.5 mM GSSG followed by the addition of Cu; Condition 4 indicates the removal of 1 mM GSSG followed by the addition of Cu. Condition 1/2 (well 9) indicates that the refolded protein underwent both conditions 1 and 2 (the removal of DTT and the addition of Cu both happened after this step). Arrows indicate the reduced and oxidized disulfide bands of monomeric ZfD3.2 and oligomers formed during refolding.

Table 1 Cu binding stoichiometry of ZfD3.2 without His tag.

Cu-loading condition[1]	Cu/protein ratio[2]
Condition 1: Remove 1mM DTT, then add Cu	0.070199
Condition 2: Add Cu, then remove 1mM DTT	0.263365
Condition 3: Remove 0.5mM GSSG, then add Cu	0.745719
Condition 4: Remove 1mM GSSG, then add Cu	0.717141

[1] The protocol shown in Fig. 4 was followed using the same samples shown in Figs. 8-9.

[2]The ratios were analyzed using AAS to measure Cu. "Protein" indicates total protein, of which approximately 70% is ZfD3.2 in the GSSG-treated samples and 50% is ZfD3.2 in the DTT-treated samples as determined by quantification of the bands in Fig. 8, lanes 2 and lanes 10.

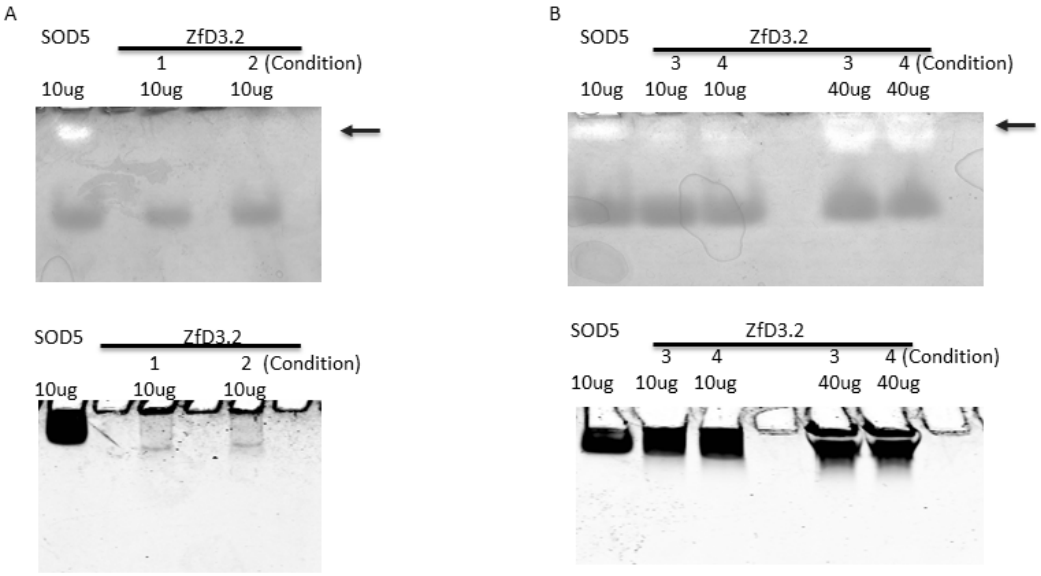

Fig. 9 ZfD3.2 SOD activity test in native gel assay.

(A) Top photo, native gel assay stained with NBT for SOD activity (see Materials and Methods) of ZfD3.2 prepared in the presence of DTT. Condition 1 indicates the removal of DTT followed by the addition of Cu; Condition 2 indicates the addition of Cu followed by the removal of DTT. Bottom photo, Coomassie stain of native gel with the same samples as in top photo. (B) Top photo, native gel stained with NBT for SOD activity of ZfD3.2 prepared in the presence of GSSG. Condition 3 indicates the removal of 0.5 mM GSSG followed by the addition of Cu; Condition 4 indicates the removal of 1 mM GSSG followed by the addition of Cu. Bottom photo, Coomassie stain of native gel with the same samples as in top photo. Arrows in both figures indicate where a white band should exist if the protein has SOD activity.

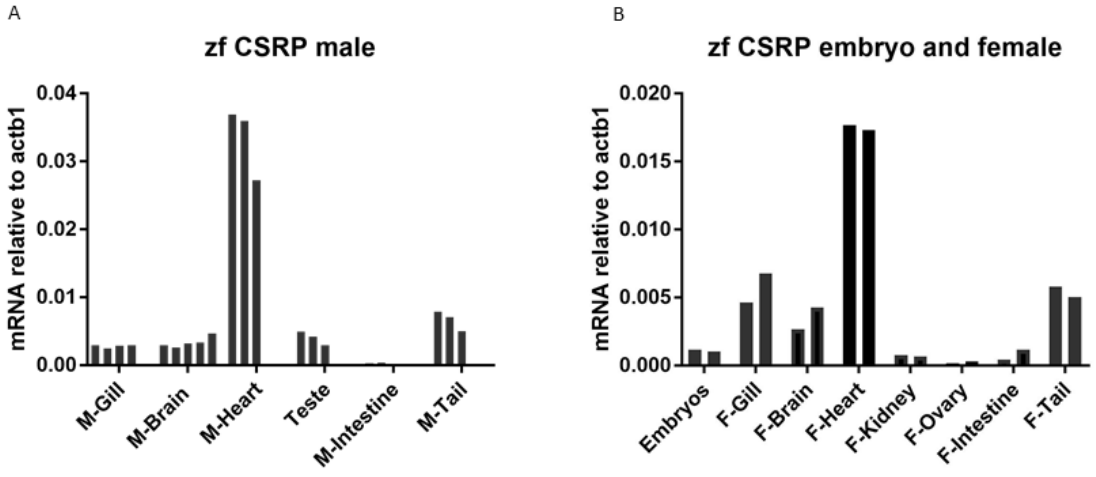

Fig. 10 qRT-PCR analysis of CSRP in zebrafish embryos and adult tissue.

CSRP mRNA expression levels in zebrafish embryos and adult female (B) or male (A) tissues. Expression levels are shown in comparison to housekeeping gene *actb1* set at the value of 1. F, female, M, male.

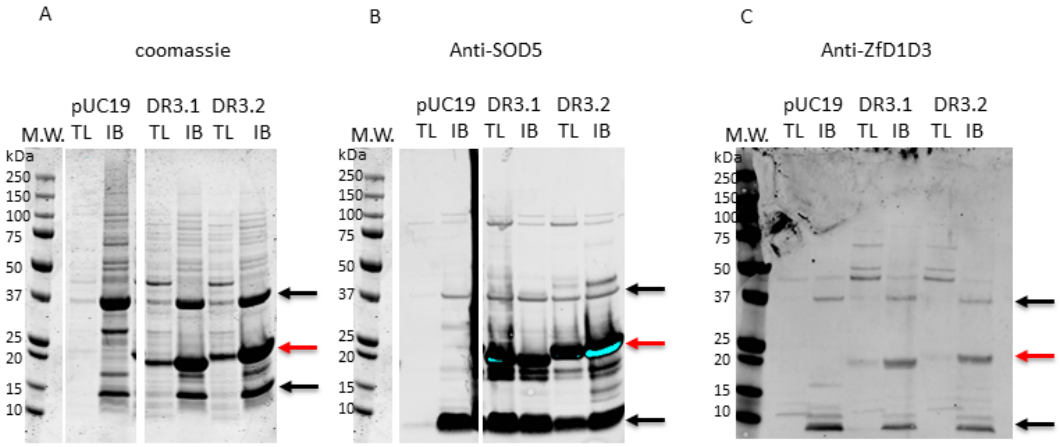

Fig. 11 Immunoblot test for anti-SOD5 versus anti-ZfD1D3 reactivity with recombinant CSRP domains.

10 µg (A) or 1 µg (B, C) of total cell lysate (TL) and the equivalent amount of inclusion body (IB) of protein from *E. coli* transformed with either pUC19 control or the indicated domains of Zf CSRP (ZfD3.1 and ZfD3.2) were run on a SDS reducing gel. The gel was either stained with Coomassie (A) or analyzed by western blot using either anti-SOD5 antibody followed by goat anti-rabbit secondary antibody (B) or anti-ZfD1D3 test bleed 2 antibody followed by goat anti-rabbit secondary antibody (C). M.W., molecular weight. Red arrows indicate the specific bands of Zf CSRP while black arrows indicate the non-specific bands of pUC19 control and Zf CSRP.

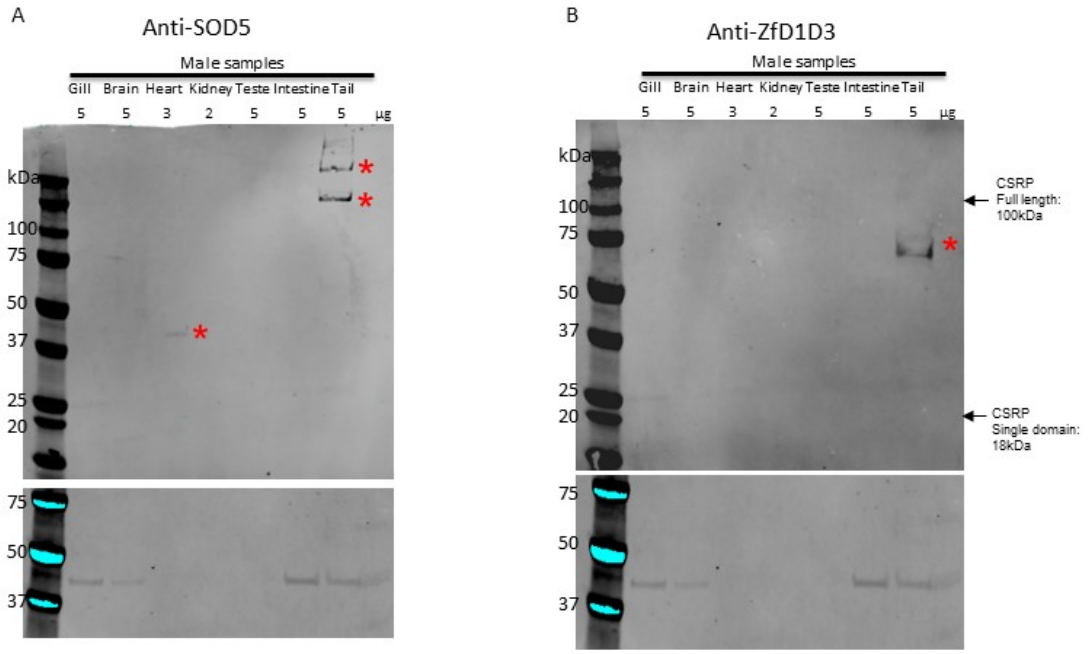

Fig. 12 Tissue specific expression of CSRP in male fish.

The indicated male fish tissues were run on a SDS reducing gel, then transferred to an iBlot2 PVDF membrane and probed using either anti-SOD5 antibody followed by goat anti-rabbit secondary antibody (A) or anti-ZfD1D3 antibody followed by goat anti-rabbit secondary antibody (B). The bottom membrane was probed using the control anti-beta actin antibody. The amount of each tissue loaded (µg) is labeled above the membranes. Arrows indicate the bands for the full-length CSRP and single CSRP domain size. Red asterisks represent the major 40, 60 and ⩾100 kDa bands see in heart and tail.

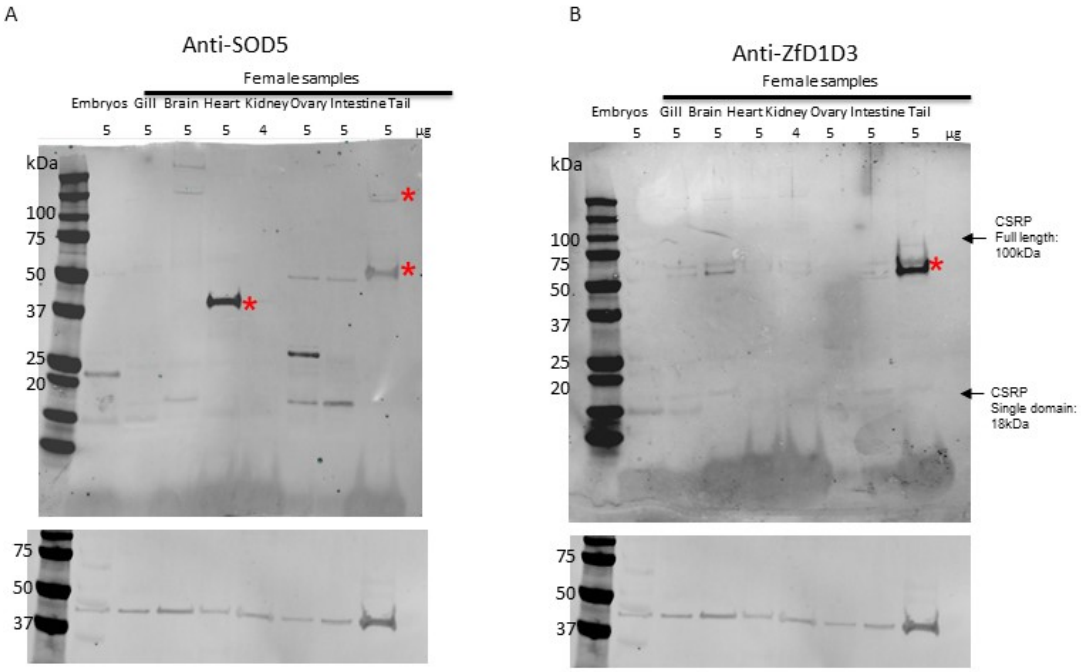

Fig. 13 Tissue specific expression of CSRP in female fish and embryos.

The indicated female fish tissues or embryos were run on the SDS reducing gel, then transferred to an iBlot2 PVDF membrane and probed using either anti-SOD5 antibody followed by goat anti-rabbit secondary antibody (A) or anti-ZfD1D3 antibody followed by goat anti-rabbit secondary antibody (B). The bottom membranes were probed with the control anti-beta actin antibody. The amount loaded (µg) for each tissue is labeled above the membranes. Arrows indicate the bands for the full length CSRP and single CSRP domain size. Red asterisks represent the major 40, 60 and ≥100 kDa bands seen in heart and tail.

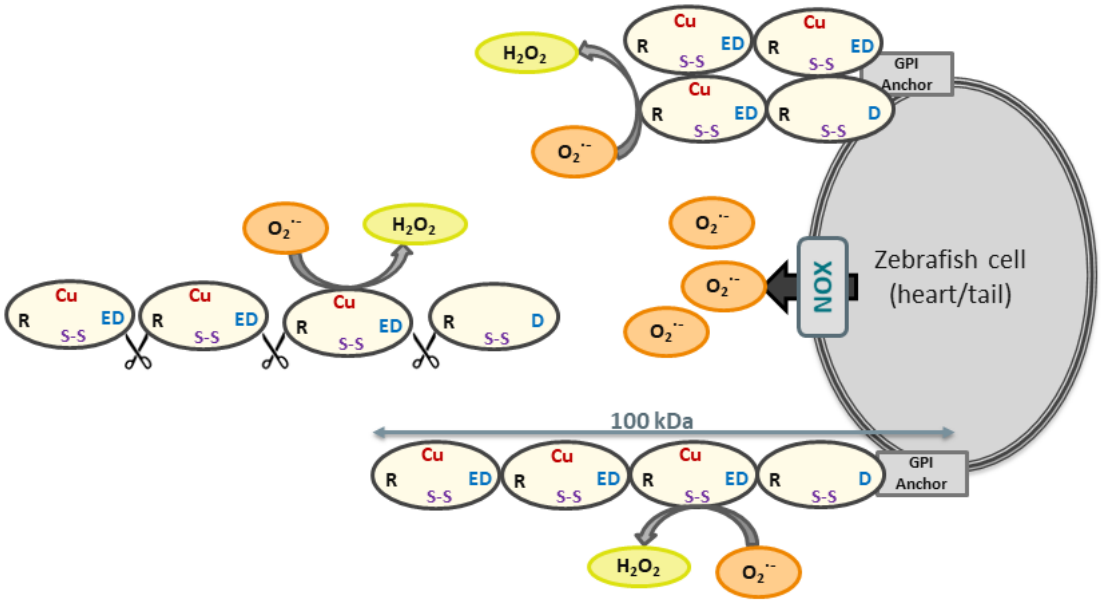

Fig. 14 Model for CSRP function.

Cartoon depicting possible expression of CSRP in zebrafish heart or tail. The CSRP of zebrafish *Danio rerio* contains four tandem domains, which may exist either as an extended polypeptide (bottom) or with interacting repeats (top). In zebrafish heart and tail tissues, we observed 40-kDa and 60-kDa species which may represent clipping of the intact polypeptide (indicated by scissors). The single CSRP domain has a SOD-like function and it is possible that in the heart or tail tissue, CSRP acts to remove extracellular superoxide produced through NOX enzymes. This could occur through either individual CSRP domains which are clipped and soluble, or the intact GPI-anchored CSRP tandem repeats. Ovals represent the Greek key β-barrel domains with active site arginine (R), Cu site, disulfide cysteines (S-S) and sequences corresponding to the active site Glu-110 and Asp-113 of *C. albicans* SOD5 (ED). The C terminal sequences for GPI anchorage are in light grey.

References

1. McCord, J.M. and I. Fridovich, *Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein).* J Biol Chem, 1969. **244**(22): p. 6049-55.
2. Carrico, R.J. and H.F. Deutsch, *The presence of zinc in human cytocuprein and some properties of the apoprotein.* J Biol Chem, 1970. **245**(4): p. 723-7.
3. Sheng, Y., et al., *Superoxide dismutases and superoxide reductases.* Chem Rev, 2014. **114**(7): p. 3854-918.
4. Fridovich, I., *Superoxide radical: an endogenous toxicant.* Annu Rev Pharmacol Toxicol, 1983. **23**: p. 239-57.
5. Cadenas, E. and K.J. Davies, *Mitochondrial free radical generation, oxidative stress, and aging.* Free Radic Biol Med, 2000. **29**(3-4): p. 222-30.
6. Juarez, J.C., et al., *Superoxide dismutase 1 (SOD1) is essential for H2O2-mediated oxidation and inactivation of phosphatases in growth factor signaling.* Proc Natl Acad Sci U S A, 2008. **105**(20): p. 7147-52.
7. Reddi, A.R. and V.C. Culotta, *SOD1 integrates signals from oxygen and glucose to repress respiration.* Cell, 2013. **152**(1-2): p. 224-35.
8. Fattman, C.L., L.M. Schaefer, and T.D. Oury, *Extracellular superoxide dismutase in biology and medicine.* Free Radic Biol Med, 2003. **35**(3): p. 236-56.
9. Broxton, C.N. and V.C. Culotta, *SOD Enzymes and Microbial Pathogens: Surviving the Oxidative Storm of Infection.* PLoS Pathog, 2016. **12**(1): p. e1005295.
10. Fenlon, L.A. and J.M. Slauch, *Phagocyte roulette in Salmonella killing.* Cell Host Microbe, 2014. **15**(1): p. 7-8.
11. Bermingham-McDonogh, O., E.B. Gralla, and J.S. Valentine, *The copper, zinc-superoxide dismutase gene of Saccharomyces cerevisiae: cloning, sequencing, and biological activity.* Proc Natl Acad Sci U S A, 1988. **85**(13): p. 4789-93.
12. Goscin, S.A. and I. Fridovich, *The purification and properties of superoxide dismutase from Saccharomyces cerevisiae.* Biochim Biophys Acta, 1972. **289**(2): p. 276-83.
13. Broxton, C.N., et al., *A role for Candida albicans superoxide dismutase enzymes in glucose signaling.* Biochem Biophys Res Commun, 2018. **495**(1): p. 814-820.
14. Crapo, J.D., et al., *Copper,zinc superoxide dismutase is primarily a cytosolic protein in human cells.* Proc Natl Acad Sci U S A, 1992. **89**(21): p. 10405-9.
15. Jaarsma, D., et al., *CuZn superoxide dismutase (SOD1) accumulates in vacuolated mitochondria in transgenic mice expressing amyotrophic lateral sclerosis-linked SOD1 mutations.* Acta Neuropathol, 2001. **102**(4): p. 293-305.
16. Okado-Matsumoto, A. and I. Fridovich, *Subcellular distribution of superoxide dismutases (SOD) in rat liver: Cu,Zn-SOD in mitochondria.* J Biol Chem, 2001. **276**(42): p. 38388-93.
17. Sturtz, L.A., et al., *A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria. A physiological role for SOD1 in guarding against mitochondrial oxidative damage.* J Biol Chem, 2001. **276**(41): p. 38084-9.
18. Turner, B.J., et al., *Impaired extracellular secretion of mutant superoxide dismutase 1 associates with neurotoxicity in familial amyotrophic lateral sclerosis.* J Neurosci, 2005. **25**(1): p. 108-17.
19. Mondola, P., et al., *The Cu, Zn Superoxide Dismutase: Not Only a Dismutase Enzyme.* Front Physiol, 2016. **7**: p. 594.
20. Tsang, C.K., et al., *Superoxide dismutase 1 acts as a nuclear transcription factor to regulate*

oxidative stress resistance. Nat Commun, 2014. **5**: p. 3446.

21. Wood, L.K. and D.J. Thiele, *Transcriptional activation in yeast in response to copper deficiency involves copper-zinc superoxide dismutase.* J Biol Chem, 2009. **284**(1): p. 404-13.
22. Puget, K. and A.M. Michelson, *Isolation of a new copper-containing superoxide dismutase bacteriocuprein.* Biochem Biophys Res Commun, 1974. **58**(3): p. 830-8.
23. Benov, L., et al., *Copper, zinc superoxide dismutase in Escherichia coli: periplasmic localization.* Arch Biochem Biophys, 1995. **319**(2): p. 508-11.
24. Imlay, K.R. and J.A. Imlay, *Cloning and analysis of sodC, encoding the copper-zinc superoxide dismutase of Escherichia coli.* J Bacteriol, 1996. **178**(9): p. 2564-71.
25. Marklund, S.L., *Human copper-containing superoxide dismutase of high molecular weight.* Proc Natl Acad Sci U S A, 1982. **79**(24): p. 7634-8.
26. Marklund, S.L., E. Holme, and L. Hellner, *Superoxide dismutase in extracellular fluids.* Clin Chim Acta, 1982. **126**(1): p. 41-51.
27. Oury, T.D., et al., *Human extracellular superoxide dismutase is a tetramer composed of two disulphide-linked dimers: a simplified, high-yield purification of extracellular superoxide dismutase.* Biochem J, 1996. **317 (Pt 1)**: p. 51-7.
28. Stromqvist, M., *Characterization of recombinant human extracellular superoxide dismutase.* J Chromatogr, 1993. **621**(2): p. 139-48.
29. Lambeth, J.D., *NOX enzymes and the biology of reactive oxygen.* Nat Rev Immunol, 2004. **4**(3): p. 181-9.
30. Battistoni, A., et al., *Crystallization and preliminary X-ray analysis of the monomeric Cu,Zn superoxide dismutase from Escherichia coli.* Protein Sci, 1996. **5**(10): p. 2125-7.
31. Tainer, J.A., et al., *Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase.* J Mol Biol, 1982. **160**(2): p. 181-217.
32. Robinett, N.G., R.L. Peterson, and V.C. Culotta, *Eukaryotic copper-only superoxide dismutases (SODs): A new class of SOD enzymes and SOD-like protein domains.* J Biol Chem, 2018. **293**(13): p. 4636-4643.
33. Hart, P.J., et al., *A structure-based mechanism for copper-zinc superoxide dismutase.* Biochemistry, 1999. **38**(7): p. 2167-78.
34. Banci, L., et al., *A characterization of copper/zinc superoxide dismutase mutants at position 124. Zinc-deficient proteins.* Eur J Biochem, 1991. **196**(1): p. 123-8.
35. Roberts, B.R., et al., *Structural characterization of zinc-deficient human superoxide dismutase and implications for ALS.* J Mol Biol, 2007. **373**(4): p. 877-90.
36. Potter, S.Z., et al., *Binding of a single zinc ion to one subunit of copper-zinc superoxide dismutase apoprotein substantially influences the structure and stability of the entire homodimeric protein.* J Am Chem Soc, 2007. **129**(15): p. 4575-83.
37. Seetharaman, S.V., et al., *Disrupted zinc-binding sites in structures of pathogenic SOD1 variants D124V and H80R.* Biochemistry, 2010. **49**(27): p. 5714-25.
38. Perry, J.J., et al., *The structural biochemistry of the superoxide dismutases.* Biochim Biophys Acta, 2010. **1804**(2): p. 245-62.
39. Fisher, C.L., et al., *The role of arginine 143 in the electrostatics and mechanism of Cu,Zn superoxide dismutase: computational and experimental evaluation by mutational analysis.* Proteins, 1994. **19**(1): p. 24-34.
40. Antonyuk, S.V., et al., *The structure of human extracellular copper-zinc superoxide dismutase*

at 1.7 A resolution: insights into heparin and collagen binding. J Mol Biol, 2009. **388**(2): p. 310-26.

41. Getzoff, E.D., et al., *Faster superoxide dismutase mutants designed by enhancing electrostatic guidance.* Nature, 1992. **358**(6384): p. 347-51.
42. Getzoff, E.D., et al., *Electrostatic recognition between superoxide and copper, zinc superoxide dismutase.* Nature, 1983. **306**(5940): p. 287-90.
43. Polticelli, F., et al., *Role of the electrostatic loop charged residues in Cu,Zn superoxide dismutase.* Protein Sci, 1998. **7**(11): p. 2354-8.
44. Martchenko, M., et al., *Superoxide dismutases in Candida albicans: transcriptional regulation and functional characterization of the hyphal-induced SOD5 gene.* Mol Biol Cell, 2004. **15**(2): p. 456-67.
45. De Groot, P.W., K.J. Hellingwerf, and F.M. Klis, *Genome-wide identification of fungal GPI proteins.* Yeast, 2003. **20**(9): p. 781-96.
46. Amorim-Vaz, S., et al., *RNA Enrichment Method for Quantitative Transcriptional Analysis of Pathogens In Vivo Applied to the Fungus Candida albicans.* MBio, 2015. **6**(5): p. e00942-15.
47. Bruno, V.M., et al., *Transcriptomic analysis of vulvovaginal candidiasis identifies a role for the NLRP3 inflammasome.* MBio, 2015. **6**(2).
48. Fanning, S., et al., *Divergent targets of Candida albicans biofilm regulator Bcr1 in vitro and in vivo.* Eukaryot Cell, 2012. **11**(7): p. 896-904.
49. Frohner, I.E., et al., *Candida albicans cell surface superoxide dismutases degrade host-derived reactive oxygen species to escape innate immune surveillance.* Mol Microbiol, 2009. **71**(1): p. 240-52.
50. Pierce, J.V., et al., *Normal adaptation of Candida albicans to the murine gastrointestinal tract requires Efg1p-dependent regulation of metabolic and host defense genes.* Eukaryot Cell, 2013. **12**(1): p. 37-49.
51. Thewes, S., et al., *In vivo and ex vivo comparative transcriptional profiling of invasive and non-invasive Candida albicans isolates identifies genes associated with tissue invasion.* Mol Microbiol, 2007. **63**(6): p. 1606-28.
52. Nantel, A., et al., *Transcription profiling of Candida albicans cells undergoing the yeast-to-hyphal transition.* Mol Biol Cell, 2002. **13**(10): p. 3452-65.
53. Gleason, J.E., et al., *Candida albicans SOD5 represents the prototype of an unprecedented class of Cu-only superoxide dismutases required for pathogen defense.* Proc Natl Acad Sci U S A, 2014. **111**(16): p. 5866-71.
54. Peterson, R.L., et al., *The Phylogeny and Active Site Design of Eukaryotic Copper-only Superoxide Dismutases.* J Biol Chem, 2016. **291**(40): p. 20911-20923.
55. Valentine, J.S., et al., *pH-dependent migration of copper(II) to the vacant zinc-binding site of zinc-free bovine erythrocyte superoxide dismutase.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4245-9.
56. Spagnolo, L., et al., *Unique features of the sodC-encoded superoxide dismutase from Mycobacterium tuberculosis, a fully functional copper-containing enzyme lacking zinc in the active site.* J Biol Chem, 2004. **279**(32): p. 33447-55.
57. Petersen, S.V., et al., *The folding of human active and inactive extracellular superoxide dismutases is an intracellular event.* J Biol Chem, 2008. **283**(22): p. 15031-6.
58. Qin, Z., et al., *Essential role for the Menkes ATPase in activation of extracellular superoxide*

dismutase: implication for vascular oxidative stress. FASEB J, 2006. **20**(2): p. 334-6.

59. Richards, T.A., et al., *Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes.* Proc Natl Acad Sci U S A, 2011. **108**(37): p. 15258-63.
60. Savory, F., G. Leonard, and T.A. Richards, *The role of horizontal gene transfer in the evolution of the oomycetes.* PLoS Pathog, 2015. **11**(5): p. e1004805.
61. Pfefferli, C. and A. Jazwinska, *The careg element reveals a common regulation of regeneration in the zebrafish myocardium and fin.* Nat Commun, 2017. **8**: p. 15151.
62. Poss, K.D., L.G. Wilson, and M.T. Keating, *Heart regeneration in zebrafish.* Science, 2002. **298**(5601): p. 2188-90.
63. Foglia, M.J. and K.D. Poss, *Building and re-building the heart by cardiomyocyte proliferation.* Development, 2016. **143**(5): p. 729-40.
64. Rieger, S. and A. Sagasti, *Hydrogen peroxide promotes injury-induced peripheral sensory axon regeneration in the zebrafish skin.* PLoS Biol, 2011. **9**(5): p. e1000621.
65. Landis, G.N. and J. Tower, *Superoxide dismutase evolution and life span regulation.* Mech Ageing Dev, 2005. **126**(3): p. 365-79.
66. Zhang, G., et al., *The oyster genome reveals stress adaptation and complexity of shell formation.* Nature, 2012. **490**(7418): p. 49-54.
67. Kobayashi, Y., et al., *Comparative analysis of seven types of superoxide dismutases for their ability to respond to oxidative stress in Bombyx mori.* Sci Rep, 2019. **9**(1): p. 2170.
68. Ringrose, J.H., et al., *Deep proteome profiling of Trichoplax adhaerens reveals remarkable features at the origin of metazoan multicellularity.* Nat Commun, 2013. **4**: p. 1408.
69. Lele, Z. and P.H. Krone, *The zebrafish as a model system in developmental, toxicological and transgenic research.* Biotechnol Adv, 1996. **14**(1): p. 57-72.
70. Spitsbergen, J.M. and M.L. Kent, *The state of the art of the zebrafish model for toxicology and toxicologic pathology research--advantages and current limitations.* Toxicol Pathol, 2003. **31 Suppl**: p. 62-87.
71. Grunwald, D.J. and J.S. Eisen, *Headwaters of the zebrafish -- emergence of a new model vertebrate.* Nat Rev Genet, 2002. **3**(9): p. 717-24.
72. Sprague, J., et al., *The Zebrafish Information Network (ZFIN): the zebrafish model organism database.* Nucleic Acids Res, 2003. **31**(1): p. 241-3.
73. Sarasamma, S., et al., *Zebrafish: A Premier Vertebrate Model for Biomedical Research in Indian Scenario.* Zebrafish, 2017. **14**(6): p. 589-605.
74. Nowik, N., et al., *Zebrafish: an animal model for research in veterinary medicine.* Pol J Vet Sci, 2015. **18**(3): p. 663-74.
75. Peterson, R.L., et al., *The Phylogeny and Active Site Design of Eukaryotic Copper-only Superoxide Dismutases.* Journal of Biological Chemistry, 2016. **291**(40): p. 20911-20923.
76. Robinett, N.G., et al., *Exploiting the vulnerable active site of a copper-only superoxide dismutase to disrupt fungal pathogenesis.* J Biol Chem, 2019. **294**(8): p. 2700-2713.
77. Ida, M., et al., *Structural basis of Cu, Zn-superoxide dismutase amyloid fibril formation involves interaction of multiple peptide core regions.* J Biochem, 2016. **159**(2): p. 247-60.
78. Suliman, H.B., M. Ali, and C.A. Piantadosi, *Superoxide dismutase-3 promotes full expression of the EPO response to hypoxia.* Blood, 2004. **104**(1): p. 43-50.
79. Hayward, L.J., et al., *Decreased metallation and activity in subsets of mutant superoxide dismutases associated with familial amyotrophic lateral sclerosis.* Journal of Biological

Chemistry, 2002. **277**(18): p. 15923-15931.

80. Gauron, C., et al., *Sustained production of ROS triggers compensatory proliferation and is required for regeneration to proceed.* Sci Rep, 2013. **3**: p. 2084.
81. Niethammer, P., et al., *A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish.* Nature, 2009. **459**(7249): p. 996-9.
82. Han, P., et al., *Hydrogen peroxide primes heart regeneration with a derepression mechanism.* Cell Res, 2014. **24**(9): p. 1091-107.
83. Chen, C.H., et al., *Multicolor Cell Barcoding Technology for Long-Term Surveillance of Epithelial Regeneration in Zebrafish.* Dev Cell, 2016. **36**(6): p. 668-80.
84. Razaghi, B., et al., *hace1 Influences zebrafish cardiac development via ROS-dependent mechanisms.* Dev Dyn, 2018. **247**(2): p. 289-303.
85. Daugaard, M., et al., *Hace1 controls ROS generation of vertebrate Rac1-dependent NADPH oxidase complexes.* Nat Commun, 2013. **4**: p. 2180.
86. Sehring, I.M., C. Jahn, and G. Weidinger, *Zebrafish fin and heart: what's special about regeneration?* Curr Opin Genet Dev, 2016. **40**: p. 48-56.
87. Yoo, S.K., et al., *Lyn is a redox sensor that mediates leukocyte wound attraction in vivo.* Nature, 2011. **480**(7375): p. 109-12.
88. Enghild, J.J., et al., *The heparin-binding domain of extracellular superoxide dismutase is proteolytically processed intracellularly during biosynthesis.* J Biol Chem, 1999. **274**(21): p. 14818-22.
89. Behrendt, M., J. Polaina, and H.Y. Naim, *Structural hierarchy of regulatory elements in the folding and transport of an intestinal multidomain protein.* J Biol Chem, 2010. **285**(6): p. 4143-52.

Mengrui Wang
615N. Wolfe Street Baltimore, MD 21205
mwang107@jh.edu | (443)722-0171

EDUCATION

Estimated 2019 **Master of Science (ScM)**
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
Project Topic: The Cu-only SOD repeat protein of the zebrafish *Danio rerio*

Honors: 2018–2019 Centennial Scholars Fund

Aug 2017- **Master of Health Science (MHS)**
May 2018 Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
Concentration: Biochemistry and Molecular Biology

Aug 2013 - **Bachelor of Engineering in Biological Engineering**
Jul 2017 East China University of Science and Technology, Shanghai, China

Aug 2013 - **Certification in Finance**
May 2017 East China University of Science and Technology, Shanghai, China

Honors: Shanghai Scholarships;
School scholarship for the integrated course;
Prize for social and community work;
Outstanding student leader;
"Shan Xing 100"Star volunteers

Jul 2014 - **Summer Exchange**
Aug 2014 University of Dundee, Scotland, UK
Concentration: Arts, Physical and Mechanical Engineering

RESEARCH TRAINING

Estimated 2019 **ScM student**
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
PI: Dr. Valeria Culotta
Project Topic: Cu-SOD Repeat Proteins (CSRP)

- Express and purify CSRP domains for protein characterization
- Analyze CSRP mRNA and protein levels in zebrafish tissues
- Compare the localization of CSRP with EC-SOD and NOX enzymes

Jan 2018-
Mar 2018

MHS Student
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
PI: Dr. Scott Bailey
Description: Basic researches in the biomedical sciences through work under the guidance of BMB faculty members

- PCR Amplification of CSM-6
- LIC cloning CSM-6 (PSAT-L as vector)
- Protein expression and purification

Mar 2018-
May 2018

MHS Student
Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
PI: Dr. Jiou Wang
Description: Basic researches in the biomedical sciences through work under the guidance of BMB faculty members

- Murine transgenic experiments, like tagging and tailing mice for genotyping
- Murine surgeries, like fresh tissue harvesting and perfusion

Oct 2015 -
May 2017

Research Assistant
University Student Innovation Experiment Program, Shanghai, China
Topic: Screening and Cloning Statin's Side Chain Precursor CTeHP oxidase

- Planned experiments to screen efficient CTeHP oxidase
- Identified relevant gene sequence that controls the expression of CTeHP oxidase
- Cloned the gene

Jul 2016 -
Aug 2016

Laboratory Assistant
Institute of Biophysics, Chinese Academy of Science, Beijing, China
PI: Dr. Lijun Bi

- Conducted basic experiment such as protein purification, RNA extraction and Western Blot for the study of tuberculosis

Jun 2015 -
Mar 2016

Program Intern
Undergraduate Students Research Program, Shanghai, China
Topic: Explore and Efficient Expression of Amylase in Low Temperature

- Selected starting strain for G. Pannorum
- Conducted experiments to clone the new α-amylase expressed in low-temperature